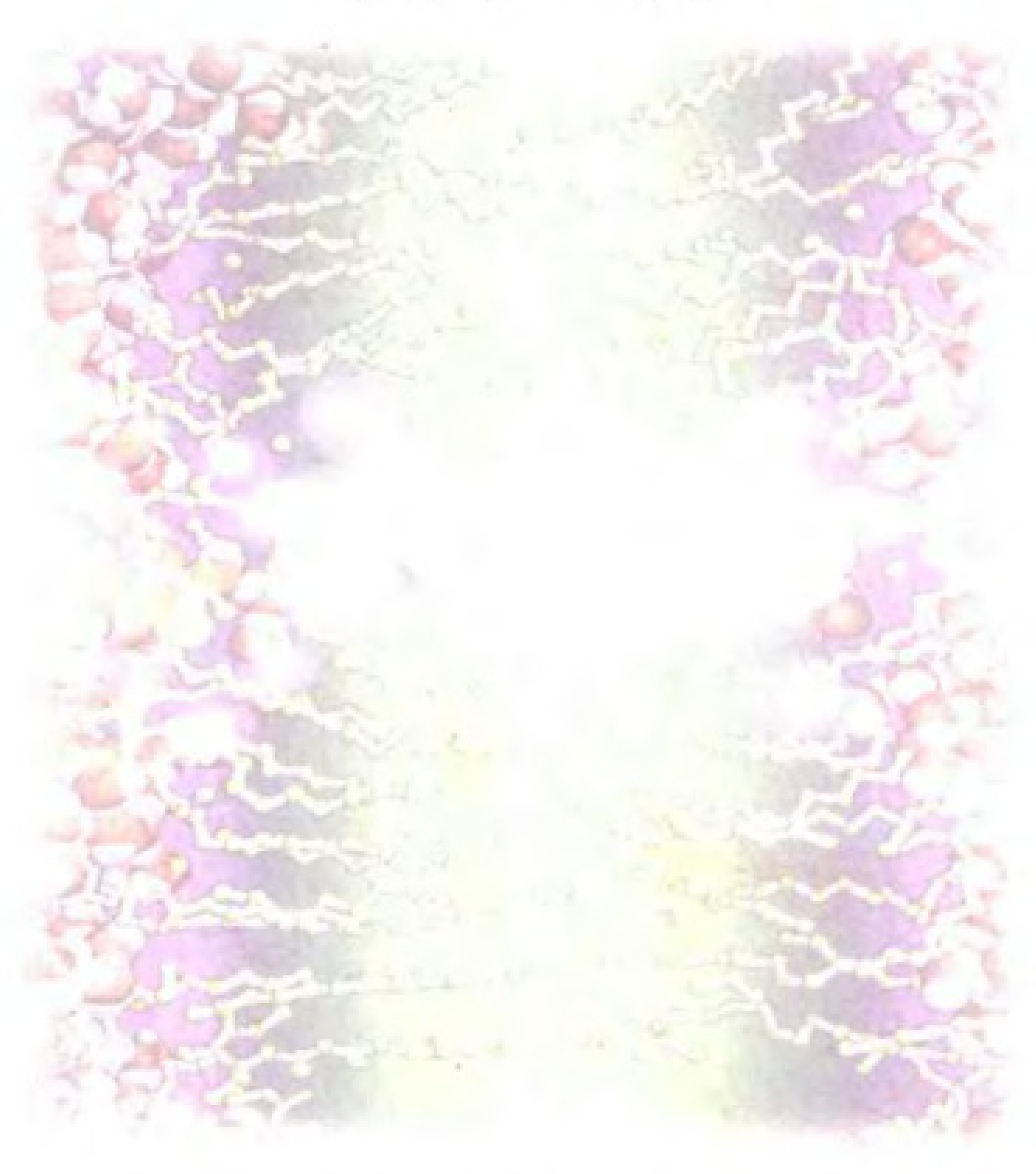
FAYI
DNA
FENXI

法医DNA分析

郑秀芬 编著



中国人民公安大学出版社

法医 DNA 分析

郑秀芬 编著

中国人民公安大学出版社
·北 京·

图书在版编目(CIP)数据

法医 DNA 分析/郑秀芬编著 .—北京:中国人民公安大学出版社,2002.4
ISBN 7-81059-965-8

Ⅰ.法… Ⅱ.郑… Ⅲ.脱氧核糖核酸-法医学鉴定 Ⅳ.D919.2

中国版本图书馆 CIP 数据核字(2002)第 016108 号

法医 DNA 分析

FAYI DNA FENXI

郑秀芬 编著

出版发行:中国人民公安大学出版社
地　　址:北京市西城区木樨地南里
邮政编码:100038
经　　销:新华书店
印　　刷:北京市振兴印刷厂印刷

版　　次:2002 年 5 月第 1 版
印　　次:2002 年 5 月第 1 次
印　　张:31.5
开　　本:787mm×1092mm　1/16
字　　数:744 千字
印　　数:0001~1000 册

ISBN 7-81059-965-8/D·818
定　　价:63 元

本社图书出现印装质量问题,由发行部负责调换
联系电话:(010)83905728

E-mail:cpep@public.bta.net.cn

前　言

法医DNA分析，通称DNA鉴定。因为它的极高的识别率和灵敏度，自建立以来的短短17年间得到了飞速的发展和广泛的应用，今天它已经成为各国刑事侦查和司法审判获取证据的重要的不可缺少的科学技术手段。

本人十多年一直从事法医DNA分析应用的研究、办案和新技术培训，一方面深感法医DNA分析技术发展非常之快，可谓“日新月异”；另一方面也深感法医DNA分析涉及遗传学、分子遗传学、群体遗传学、生物化学、分子生物学、生物统计学和计算机学等众多学科，是多科的交叉、综合，而目前既全面系统又以新技术为主，既有简单必要的基础理论又有详尽的实验操作的法医DNA分析的书却比较少。有鉴于此，结合自己多年的研究成果、实际检案和培训经验，参考国内外有关文献著作，编写了本书，希望为我国法医DNA分析发展作出一些绵薄之力。

本人在此书中力求对子法医DNA分析技术从理论基础、技术操作到结果应用有一个全面的总体的介绍，充分和概括地反映这一领域的研究应用成果，对于当前的主流应用技术给予详尽的叙述，对子其发展有一定的前瞻性展示。对涉及的基础理论和技术给予必要的笔墨。在章节内容安排上逐步展开，以便于阅读。力求做到内容实用、先进、简明易读。

在本书编写过程中得到了李伯龄主任法医师、倪锦堂主任法医师、杨庆恩教授、侯一平教授、叶健主任法医师等的帮助和鼓励，对书稿提出了许多宝贵意见，在此表示衷心的感谢。

法医DNA分析技术发展很快，编者水平有限，错漏之处在所难免，欢迎批评指正。

作者　郑秀芬

2002年1月

序　一

很高兴应邀为本书作序。法医 DNA 技术一直是我多年以来在专业上所特别关注的领域。

20 世纪最后 2 个年代在刑事科学领域里最大的进展莫过于 DNA 鉴定技术的出现和应用。它为成千上万的各种案件的侦破和解决提供了强有力的科学证据。

DNA 鉴定技术由于是从生物的分子—基因水平揭露生物的特性，它的极高鉴定能力和比较广泛应用的可能性以及在侦审中的独特作用，使其一直受到高度的重视，倍受关注，分子生物和基因等基础和相关领域的进展使其发展飞快。自 1985 年英国遗传学家 Alec Jeffreys 建立了 DNA 指纹技术以来，新技术新方法不断出现，可应用的各种遗传标记与日俱增，所探讨解决的问题也在逐步扩大。今天 DNA 技术已成为个人识别和血缘鉴定的最有效方法，并且正在朝着使这些鉴定向可以检验各种条件下的检材和更进一步准确、微量、快速、自动化及数据库的方向发展。与此同时，正在全方位的研究开发和利用 DNA 所具有可揭示生物各种特征信息的潜能，并从人扩大到动物，乃至其他生物，为刑侦和司法鉴定服务。可以说 DNA 鉴定技术的发展前途广阔，方兴未艾。

本书的作者是一位十多年来一直在从事法医 DNA 分析应用的研究、办案和新技术培训的年轻专业工作者。书中对于法医 DNA 分析这一涉及众多学科的新技术，做了比较全面的系统的既有简明的理论基础又有详尽的实验操作技术的介绍。广泛地吸收了国内外他人的学术成就，融人了作者自己的科研成果和办案经验。书中对于人以外的其他生物的 DNA 鉴定，对于 DNA 芯片和单核苷酸多态性等 DNA 分析应用的新内容和新技术，也有较多涉及。读后感到这是一本内容丰富、新信息多、实用性强、有前瞻性的好书。无论是对于从事法医 DNA 分析的工作者或者想要了解 DNA 鉴定技术的刑侦、司法审判、律师或其他读者都将会有很大裨益。

借为本书作序之机，希望能不断地有反映法医 DNA 分析发展的新书问世，以促进 DNA 技术的应用和发展。

李伯龄

2002 年 3 月

序 二

20世纪70年代，法医物证个体识别鉴定靠血型。受蛋白质水平遗传标记的多态性和检测技术的限制，在个体识别鉴定，只能是“排除是肯定的，认定是相对的”。曾经梦想只用一种高多态性遗传标记系统能够区别世界上所有的个体，梦想只用一种简单的方法能够在最短的时间内作出鉴定结论。

划时代的技术突破出现在1985年，英国遗传学家Jeffreys建立DNA指纹技术。应用人肌红蛋白基因内含子小卫星探针33.6和33.15制作的RFLP图，具有惊人的个体特异性，能够区分世界上所有的个体。DNA指纹实现了法医学的第一个梦想，DNA指纹广泛应用在各类案件调查中。有关卫星DNA－VNTR分型的理论研究逐渐深人，法医学研究的目光开始集中在DNA片段长度遗传标记，认识到卫星DNA－VNTR序列分型将是法医物证鉴定的希望所在。

法医物证学第二个梦想的实现靠Mullis的PCR技术。PCR近乎一种魔法，方法简单得令人不可思议：把模板DNA、dNTP、DNA聚合酶、引物同置一个反应管内，通过三个温度点的循环，数小时内，靶DNA片段被扩增近100万倍。直接采用PCR扩增VNTR基因座的片段长度等位基因，电泳检测基因型。利于高效、简单的PCR和高多态性特征的VN-TR遗传标记，达到了准确、快速、认定同一性的目标，PCR－STR分型已经成为目前个体识别鉴定的主流技术。近十年，法医物证鉴定技术连跨两大步，是进展最快的时期，尤以PCR衍生的各类新技术，新的DNA标记系统不断发现，法医物证鉴定技术研究突飞猛进。荧光标记STR复合扩增检测实现了自动化、标准化，实现了微机联网，异地查询的目标，CODIS计划就是一个成功的范例。

2001年，人类基因组计划（HGP）完成了第一张人类基因组工作草图，单核苷酸多态性（SNP）引起法医学界的注意，SNP可能是最能反映群体和个体遗传特征的一种DNA标记。应用高通量的生物芯片技术检测多SNP座位成为目前法医学应用研究的前沿课题。这是一个新的期望，也是一种新的挑战。

在世纪交替的时刻，我国的DNA分析技术发展很快，法医学者投人了极大的热忱。早在1987年，我国的DNA指纹技术开始应用，在数以千计的物证鉴定中积累了可贵的实践经验，在侦破大、要案中起到了关键作用。进人20世纪90年代，PCR检测技术开始在国内迅速推广，PCR－STR分型技术普及到了基层。在打击犯罪的复杂斗争中，DNA分析的作用越来越重要。各地司法机关投人大量的人力、物力，引进自动化多STR基因座分型检测仪器和技术，我国的DNA分析已基本与国际接轨。值得注意的是，在DNA分析技术普及的10余年中，培养出一批青年技术队伍，他们思想素质好，工作热情高，基础扎实，业务精通，活跃在与犯罪作斗争的第一线。他们是我国法医事业的希望。

作者郑秀芬，十多年来，通过一线工作的锻炼，积累了丰富的实践经验。法医DNA分析涉及医学、遗传学、分子生物学、生物统计学、免疫学等多种学科理论。由于学科进展很快，新的理论、技术层出不穷，让人目不暇接。作者力求从DNA分析的基础理论人

手，概括了DNA指纹、PCR-STR分型、DNA序列测定和法医统计学等主要内容。全书由浅入深，突出理论与实践的结合的原则，结合国内DNA分析的现状和各类物证材料特点，总结出许多宝贵的实践经验。书中还介绍了近年来许多新的技术资料，对DNA分析的技术前景作了展望。从办案第一线工作的角度来审视这本《法医DNA分析》，将会有更深刻的体会，更具特殊的实用价值。

杨庆恩

2002年3月



作者简介

郑秀芬，女，浙江余姚人，中国农业大学理科硕士。毕业后分配到公安部第二研究所工作至今,一直从事法医DNA分析的科研、办案和新技术培训工作。副主任法医师、兼中国人民公安大学硕士研究生导师。二次获国家科技进步二等奖，多次获公安部科技进步奖。发表学术论文50余篇，参加编书3部。1998年～1999年任日本秋田大学客座研究员。

目　录

第一章 绪 论

1985年英国遗传学家 Alec Jeffreys 建立了 DNA 指纹技术，并用它成功地进行第一起移民案涉及的亲子鉴定，开辟了法医物证 DNA 分析的先河，在个体同一认定和亲子鉴定上实现了从否定、排除到认定的质的飞跃，使生物物证鉴定进入了新的纪元，为刑事侦查、司法审判提供了有力的证据。

法医 DNA 技术在此后短短的10多年里得到了飞速的发展和广泛的应用。我国自开展法医 DNA 分析以来，应用该技术成功地鉴定了大量疑难案件物证，在众多的刑事案件的侦破、民事案件的解决和一些重大事故的尸源认定上发挥了重要作用。在2000年公安部开展的打击拐卖妇女儿童专项斗争中，应用 DNA 分析技术进行了大量的亲子鉴定，辨明了被拐儿童的身份，使他们回到了亲生父母身边。所有这些使广大公安政法人员以及百姓逐渐了解、认识法医 DNA 技术的重要作用。媒体对一些重大案件的报道，特别是对美国辛普生谋杀前妻和前妻男友案的世纪审判等的广泛报道和炒作，使法医 DNA 技术得到更广泛的宣传，已经是家喻户晓，老少皆知。

今天随着法制建设的不断完善和 DNA 技术的发展，要求进行 DNA 检验的案件越来越多，法医 DNA 分析如今已成为生物物证检验的常规技术。

1.1 法医 DNA 分析的研究内容

法医 DNA 分析，是应用现代 DNA 技术分析 DNA 遗传标记在群体中的分布与传递规律，确定分析样品的一致性与遗传关系，为侦查破案和司法审判提供证据的一门科学技术。法医 DNA 分析又称法医分子生物学，或法医分子遗传学，通称 DNA 鉴定，涉及遗传学、分子遗传学、群体遗传学、生物化学、分子生物学、生物统计学和计算机学等学科，是多门学科的交叉和综合应用。

DNA（deoxyribonucleic acid）是脱氧核糖核酸的简称。它主要由四种碱基 A、T、G 和 C 组成，是存在于细胞中的遗传物质。遗传物质包含了任一机体发育和功能所必需的全部信息。除同卵双生外，每个生物包括人的这四种碱基排列顺序——DNA 碱基序列是不同的，独一无二的。人类 DNA 含有 3.2×10^9 个碱基对，其中约99.9%的 DNA 序列是相同的，另外的0.1%在个体间有差异。这些 DNA 序列上的差异有的在个人的特征如眼睛、发色和肤色等上表现出来，更多的是不表现在个人的生理外观特征上，必须用实验室的特殊技术才能测定出来，其他生物也是如此。个体间的 DNA 差异及其遗传规律，个体的 DNA 差异与生物性状的关系，对于解决侦查破案、司法审判遇到的问题非常重要，这些问题是法庭科学家们所关心的和所要研究的。

自1985年 DNA 指纹技术诞生以来，经过十多年的发展，法医 DNA 分析已建立了 DNA 指纹技术、PCR 扩增片段长度多态性分析技术和线粒体 DNA 测序等三大主要技术，并在它们基础上发展出一些新技术方法，如 MVR - PCR，PCR - SSOP，SSP - PCR 等，共同形

成了法医 DNA 分析技术。这些技术方法在个体识别率、检验时间及灵敏度上有所差别（图 1-1），随着各种新的分析技术的不断建立，检验的遗传标记的不断增加，提供的信息类型不断扩大，涉及的理论不断深入和完善，各种数据的不断丰富，使法医 DNA 分析不断地得到了发展。DNA 分析技术不断提高灵敏度、加快检验速度，从原来的需要大量生物检材才能满足检验到现在可以进行单个细胞检验，检验的时间从先前的 6~8 天到目前的数个小时，不论是有核细胞，还是无核细胞，都可以应用不同的方法进行检验，大大拓宽了生物物证检验的范围，提高了生物物证的有效利用率。逐步向自动化发展，日益满足刑事侦查、司法审判的需要。

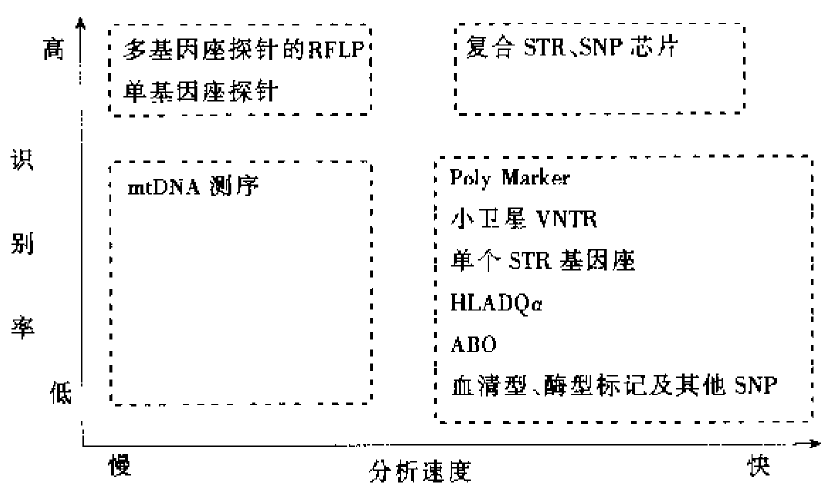


图 1-1 DNA 遗传标记与技术比较

1.2 法医 DNA 分析的对象

DNA 存在于各种生物的细胞内，因此理论上含有细胞的各种生物材料都可以进行法医 DNA 分析。从物证鉴定的角度可以分为四类：(1) 人体的各种体液（斑痕）、组织、毛发及表皮脱落物；(2) 各种动物成分；(3) 植物的根、茎、叶、花、果与种子；(4) 细菌与真菌等微生物。

人体的各种体液（斑痕）及组织是法医鉴定的各类案件中涉及最多的生物物证，对于这一部分 DNA 分析的应用研究最多，也最充分，是法医 DNA 分析的主要客体。在法医鉴定的各类案件中遇到的各种动物的毛等已成为法医 DNA 分析的检材。近年来已经开始注意应用法医 DNA 分析涉及动物引起的纠纷案件物证。另外，法医 DNA 分析开始对现场、犯罪嫌疑人等处的动植物及土壤微生物痕迹进行检验，推断嫌疑人是否去过案发现场，确定嫌疑人与案件的关系。虽然人及某些动物以外的这些生物物证的 DNA 分析应用，还要取决于对其 DNA 特征与数据的研究进展、应用开发以及侦审工作的需要，但他们势必将逐渐成为法医 DNA 分析的常规检验对象之一。

1.3 法医 DNA 分析的任务

法医 DNA 分析的最大优势就是它的高识别率，能进行个体同一认定与亲子鉴定，这也是生物物证鉴定的主要目的。因此，人以及动物的个体识别与亲子鉴定是法医 DNA 分

析的主要任务。由于DNA遗传标记非常丰富，Y染色体的遗传标记是父系遗传，线粒体DNA是母系遗传，从而可以进行几代，甚至多代的亲缘关系鉴定，使昔日根据一般遗传标记的亲子鉴定的有限范围得到较大的扩大，实际法医DNA分析的亲子鉴定已成为血缘关系鉴定。

DNA是一切生物的遗传物质，决定着生物的个体性状、特征以及生理状况等，因此从理论上说，确定个体的其他特征，也将是未来法医DNA分析可能担负的任务。

1.4 法医DNA分析在案件中的应用

法医DNA分析可以应用于所有涉及有细胞的物证的案件中。可以进行法医DNA检验的案件有：凶杀、性犯罪、恶性械斗案；在现场无意中留下指纹、毛发、唾液（烟头、口香糖等）或面罩的盗窃案、抢劫案等，以及一些投毒案等；碎尸案尸源认定；空难或其他灾难性事件中尸源认定、交通事故中尸源认定与确定带有可疑血痕、组织或毛发等的嫌疑肇事车是否为肇事车的案件；拐卖儿童、离婚、移民、入户、寻亲、遗产继承、计划生育中超生和医院错抱婴儿的亲子鉴定；器官移植、医疗纠纷中的个体识别与亲子鉴定等等，法医DNA分析的应用范围非常广泛。

1.5 法医DNA分析的特点

1. 遗传标记多，信息量大。DNA多态性标记在生物体内的分布远多于蛋白质标记，标记多态性程度大大超过蛋白质标记。仅分析其中的少部分遗传标记就完全可以直接进行个体同一认定与亲子鉴定，具有非常高的识别率和鉴定能力。

2. 标记的遗传类型全，有按孟德尔分离和自由组合律遗传的常染色体遗传标记，也有只父系遗传的Y染色体的遗传标记和母系遗传的线粒体DNA序列差异。利用DNA遗传标记可以进行几代甚至隔代、多代的血缘关系的鉴定。

3. 可分析的物证种类多，存在比较广泛。凡是含有细胞的生物材料均能进行DNA分析，这类物质可在很多种案件中提取到。具有检材的多样性与应用的广泛性。

4. 对于自然环境具有相对比较高的稳定性。DNA遗传标记相对稳定，保存几十年的血痕、精斑仍可以进行DNA分析，甚至历时数千年的木乃伊也可以获得含有一定遗传信息的结果。

5. 方法灵敏，检材用量少。当前的主流技术都是以PCR技术为基础，具有比较高的灵敏度，理论上一个细胞的DNA即可分析。实际应用中，灵敏度为ng级DNA，只需很少的检材，即可获得检验结果。

6. 同一个体物证具有比对的通用性。同一个体的各种组织DNA具有相同的序列特征，即个体组织的同一性，犹如全息摄影，因此可以用同一个体的任一部分组织或体液得到该个体的相同的全部信息，进行比对和鉴定，而不必一定要采用相同的组织。

7. 自动化程度高，速度快。

8. 可以建立数据库，进行查档比对，检索破案。

1.6 法医 DNA 分析的沿革与国内外进展

1.6.1 沿革

1944 年，Oswald Avery 证实了细胞成分 DNA（脱氧核糖核酶 deoxyribonucleic Acid）为遗传特征从上代向下一代传递的载体。

1953 年，James Watson 和 Francis Crick 阐明了 DNA 分子结构为双螺旋结构，DNA 分子本身的特征很好地解释了包括稳定地复制自身并从上一代传递给下一代能力的特性。

1980 年，David Bostein 和同事首先发现在遗传水平上人群间存在小的差异，用 RFLP 技术检测这种差异。

1984 年，Alec Jeffreys 在调查疾病的 DNA 遗传标记时，发现了 RFLP 技术在个人识别科学中的特独应用，RFLP 图谱就像人的手指纹一样，每个人不同，独一无二的，可以进行像指纹一样的个体同一认定，因此形象地称 RFLP 图谱为 DNA 指纹图（DNA fingerprint）。

1985 年，首次进行一起移民涉及的亲子鉴定，1986 年首次应用于刑事案件——二起强奸杀人案。

1985 年，Kary Mullis 发明了多聚酶链反应（polymerase chain Reaction，简称为 PCR），同时 4 色荧光自动测序问世。

1990 年，PCR 技术在法医学应用，第一个用 PCR 分析的遗传标记为 HLA DQα。

1990 年，Jeffreys 报道小卫星 D1S8 基因座串联重复单位内部存在差异，1991 年报道用 MVR - PCR 分析小卫星可变重复序列。

1991 年，荧光标记复合扩增分析 STR 基因座和应用 Chelex - 100 方法提取 DNA。

1992 年，开始报道用毛细管电泳分析 STR 结果。

1993 年，第一个商业 STR 试剂盒面世，PCR 同步扩增 Amelogenin 基因座进行性别鉴定。

1996 年，多色荧光标记 STR 试剂盒问世，开始报道用 MALDI - TOF 分析 STR 扩增产物和线粒体 DNA 基因芯片。

1997 年，开始大量报道 Y - STR 基因座。

1998 年，2 000 个 SNP 杂交芯片闻世。

1999 年，Y - STR 得到比较广泛的研究、应用。

2001 年，人类基因组计划框架初步完成。

1.6.2 国内外进展

DNA 指纹技术问世后受到了各国司法、警察和法庭科学界的高度重视，各国广泛研究和采用 DNA 技术进行办案。欧洲各国初期广泛采用的主要是 DNA 纹印系统，北美除 DNA 纹印系统外，开发和应用了 HLA DQα 和 Polymarker 等系统，现在世界各国广泛使用的是复合 STR 系统，并进行 DNA 的直接测序分析（主要用于 mtDNA）。

我国于 1986 年开始法医 DNA 分析技术应用研究，1988 年有研究成果报道，1989 年正式应用于实际案件鉴定。研究和应用的法医 DNA 分析主要有 DNA 指纹（3′HVR - α - 珠蛋白探针，Myo 探针等）、小卫星（VNTR）扩增片段长度多态性（D1S80，APOB 等 ）、微卫星（STR - PCR）复合扩增 、数字编码小卫星系统、线粒体 DNA 测序、DNA 的性别鉴定

等等。现在广泛使用的是复合 STR 系统和 DNA 的直接测序分析（主要用于 mtDNA）。

20 世纪末的微制造技术解放了集成电路工业，加快了计算机的速度和运行能力。同样微制造技术使 DNA 分析中的样品制备和一些分析步骤微型化，设计一些仪器可以在犯罪现场进行生物物证的分析，使 DNA 分析比传统实验室分析得更快、成本更低。目前在研究开发的主要有以下三个微型化 DNA 分析仪器：微芯片毛细管电泳装置（microchip capillary electrophoresis device）、微型热循环仪（miniature thermal cycler）和杂交阵列（hybridization array）。

未来的法医 DNA 分析发展方向是微型化、智能化、网络化与全自动化，从 DNA 分析的结果中获得更多的个体信息。

第二章 DNA 分型的科学基础

2.1 DNA 的分子生物学

2.1.1 一级结构

每一个生物均由细胞组成，由细胞膜包围的细胞叫真核细胞，无细胞膜的则为原核细胞，真核细胞组成的生物叫真核生物，如人、动物和植物等。真核细胞 DNA 主要存在于细胞核中，被组蛋白紧密包装在染色体中。每一个染色体含有一个 DNA 分子，染色体 DNA 分子是一个线状大分子，人的二倍体细胞含有 23 对染色体，有 23 对染色体 DNA 分子，人的一切遗传信息都包含在这 23 对分子中。除了核 DNA 外，线粒体、叶绿体（植物）中也含有自身的 DNA，它们是较小的环状分子，能自己进行复制、转录。

DNA 有时被叫做生命的蓝图。蓝图的信息被编码在 DNA 的 A（adenine，腺嘌呤）、T（thymine，胸腺嘧啶）、G（guanine，鸟嘌呤）和 C（cytosine，胞嘧啶）四个化学基本骨架中。碱基和糖连在一起叫核苷肽（nucleoside）（表 2-1），当核苷肽再与磷酸相连则称之为核苷酸（nucleotide）（图 2-1），许多个核苷酸连在一起成为多核苷酸链（polynucleotide chain），形成核酸。根据组成核酸的五碳糖是否脱氧，将核酸分为脱氧核糖核酸（deoxyribonucleic acid，缩写为 DNA）和核糖核酸（ribonucleic acid，缩写为 RNA）。

表 2-1 组成核酸的主要碱基与相应的核苷酸

碱基	核苷	核苷酸	RNA	DNA
腺嘌呤（adenine）	腺苷（adenosine）	腺苷酸（adenylic acid）	AMP	dAMP
鸟嘌呤（guanine）	鸟苷（guanosine）	鸟苷酸（guanylic acid）	GMP	dGMP
胞嘧啶（cytosine）	胞苷（cytodine）	胞苷酸（cytodylic acid）	CMP	dCMP
胸腺嘧啶（thymine）	胸苷（thymidine）	胸苷酸（thymidylic acid）		dTMP
尿嘧啶 uracil	尿苷（uridine）	尿苷酸（uridylic acid）	UMP	

这些碱基单位通过五碳糖（脱氧核糖）和磷酸，以线状方式互相连在一起，碱基就像线上的珠子一样。DNA 的结构可以分为一级结构、二级结构及三级结构。一级结构指组成核酸的诸多核苷酸之间连键的性质，以及核苷酸排列的顺序。

不同的碱基及其不同的排列序列构成不同的 DNA 序列，包含不同的遗传信息。碱基的特定序列决定着每个人的遗传特征。组成 DNA 的碱基有以下特点（即 Chargaff 规则）：

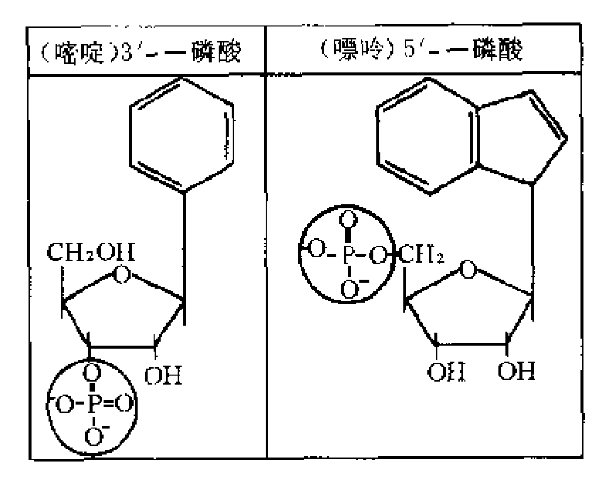
1. 所有 DNA 中腺嘌呤与胸腺嘧啶的克分子含量相等，即 A = T，鸟嘌呤与胞嘧啶的克分子含量相等，即 G = C。因此，嘌呤的总数等于嘧啶的总数，即 A + G = T + C；

2. DNA 碱基组成只有种的特异性，即不同生物种的 DNA 具有自己独特的碱基组成；

3. DNA 碱基组成没有组织器官的特异性，任何组织器官的 DNA 碱基序列均相同；

4.DNA的碱基序列不随年龄、营养状态、环境条件的不同而改变。

以上2、3、4点是法医DNA进行个人同一认定的基础。



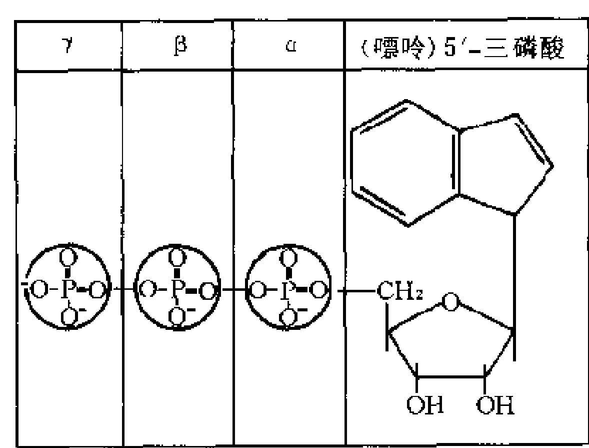


图2-1 碱基、核苷酸结构

2.1.2 DNA的双螺旋结构、碱基互补原则与理化性质

1.DNA的双螺旋结构

在自然条件下，DNA分子以双螺旋方式（二级结构）存在（图2-2），二个带状互相缠绕在一起，通过像梯子档一样的横杠紧紧地连在一起，每一横杠由二个碱基组成，这二个碱基间通过氢键连接，每一横杠互相配对的二个碱基称为碱基对，四个碱基按A与T，G与C互补配对，这种碱基互补配对（complementary base pairing）的原则叫碱基互补原则。在所有DNA分型系统中，不论是RFLP的探针设计、分子杂交，还是PCR的引物设计和链的延伸，都基于这个碱基互补原则，碱基互补原则贯穿于整个DNA分析技术。当双螺旋完整时，也就是两条链完全连在一起时，此时的DNA称为双链DNA（double strain DNA，dsDNA）。当螺旋的二半分开仅有一条时，不论在自然状态还是在试管（体外），称为单链DNA（single strain DNA，ssDNA）。两条链之间的连接力并不十分强，如100℃高温可使两条链分开，当DNA冷却下来后，二条链又可重新结合起来，这在DNA分析中十分重要。

DNA分子具有这样特殊的双螺旋结构，在理化性质与其他的分子表现不同。

2.DNA的变性与复性（熔点，退火，Tm值）

通过热或曝露于低盐浓度可以打断配对碱基间的氢键，当所有氢键被打断时，双螺旋的二条链彻底分开，这种链分开的过程叫变性（denaturation）或叫熔化（melting）。变性是DNA的一个极其重要的理化性质，它不引起DNA分子量的降低。

引起变性的因素很多，加热引起的变性叫热变性，由酸碱度改变而引起的则叫酸碱变性；乙醇、丙酮等有机溶剂也可以使DNA发生变性，导致物理性质的改变。变性后的DNA变成单链，在波长260nm处的吸收增加，称为增色效益，因此dsDNA和ssDNA的260nm处吸收不同，若将A260的增加作为温度的函数作图，可得到解链曲线或熔解曲线(melting curve)（图2-3），DNA链分离的温度范围中间点温度称为熔点温度或解链温度(melting temperature)，用Tm表示。Tm取决于分子中G-C碱基对的含量，因为G-C碱基

对中有 3 个氢键，A - T 对只有二个氢键，G - C 对含量越高，分开二条链所需的能量越多。Tm 与 GC 含量的关系是每增加 1%GC 含量，Tm 增加 0.4℃。最佳生理环境下的 DNA 溶液，Tm 一般在 85 ~ 95℃，甲酰胺等某些试剂可以破坏氢键，Tm 明显降低。变性的 DNA 在适当的条件下又可以变成双链，称之为复性（renaturation）。

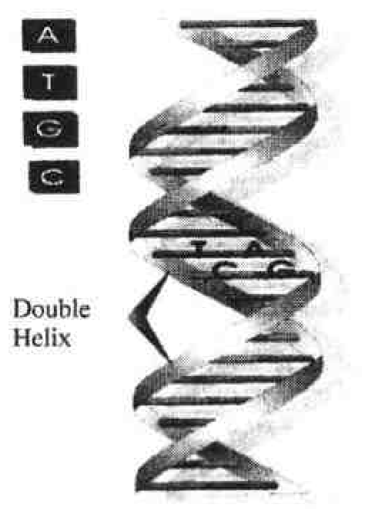


图 2 - 2　DNA 分子的二级结构

图 2 - 3　DNA 解链曲线

复性的速度受诸多因素的影响：

（1）DNA 的浓度：DNA 浓度直接影响到 DNA 单链之间的碰撞几率，DNA 浓度越大，复性速度越快；

（2）DNA 的分子大小：DNA 分子越大，扩散越慢，也难形成正确配对，复性速度慢；

（3）温度：温度过高，有利于 DNA 变性而不利于复性；温度过低，少数碱基配对形成的局部双链不易解离，难以杂交；

（4）离子强度：离子强度过低，不利于杂交，常采用 0.15 ~ 1.0mol/L 溶液进行复性；

（5）分子的复杂性：DNA 分子越复杂，复性速度越慢。

在体外，互补的短小单链 DNA 片段也表现出相互的特异亲合性，可以用已知碱基序列的探针来检测未知的序列。在适当条件下，互补的 DNA 片段可以找到互补的另一部分，重新结合在一起，在技术上称其为退火或杂交。

3. 粘度

高分子溶液比普通溶液的粘度要大得多，不规则线团分子比球形分子的粘度大，而线形分子的粘度更大。DNA 的长度与其直径之比达 10^7。即使极稀的 DNA 溶液，也有较大的粘度。当 DNA 溶液因受热或在其他因素作用下发生螺旋→线团转变时，粘度降低。所以可用粘度作为 DNA 变性的指标。

4.DNA 紫外吸收

核酸的组成成分嘌呤碱与嘧啶碱对紫外光具有强烈的吸收，所以 DNA 也有强烈的紫外吸收。其最大吸收峰在 260nm 处，而蛋白质在 280nm 处，两者不同。根据在 260nm 处吸收峰值可以测定 DNA 溶液的浓度，计算 260nm 与 280nm 处吸收峰值比可以判断 DNA 样品

是否纯净。

5. 酸碱性质

DNA 分子组成中有磷酸分子，连接二核苷酸，当溶液的 pH 值高于 4 时，DNA 呈多阴离子状态，在电场作用下带负电。

2.1.3 DNA 的半保留复制

原则上，互补碱基配对保证了 DNA 分子的准确复制，将遗传信息传递给下一代。在体内，双螺旋 DNA 分子在解旋酶、解链酶作用下，变成单链 DNA 分子，每条链作为合成新链的模板，称为引物的短小寡核苷酸分子根据碱基配对原则各自结合到互补的单链 DNA 分子中，随后在 DNA 聚合酶（polymerase）作用下，将新的骨架成分——碱基，按单链分子上的碱基互补序列连接在引物后，一个碱基接一个碱基从链的 5′→3′延长，即 DNA 的延长方向是 5′→3′。合成与旧链互补的新链，结果产生二个与原来相同的分子，传递给子代。这个过程可以在体外模拟，这就是聚合酶链式反应（polymerase chain reaction, PCR），在以后章节会详细地介绍 PCR。

2.2 人类遗传学基础

遗传学是研究生物有机体的遗传和变异的一门学科。遗传学研究不仅涉及个体，如从父母代到子女代的特征的遗传，又涉及群体，各种族个体遗传特征在群体中的变异，这二个范畴组成了在法医物证鉴定中 DNA 应用的整体。

2.2.1 遗传的物质基础

所有生物均由细胞组成，细胞是生命的最小单位，每个细胞大小约为一根头发直径的 1/10，大多数体细胞（除红细胞外）均含有一个小的完整的核，它是组成细胞组织的中心，遗传信息就存在于这个细胞的核中，以染色体的物理结构出现，通常染色体以整个完整单位从父母代传递给下代。因此，存在于同一染色体上的遗传标记也一起遗传下来。他们表现为遗传连锁（genetic linkage）。相反的，位于不同染色体上的标记一般是互相独立遗传的，这个原则叫随机分配（random assortment），表现为随机分配的遗传标记是不一起遗传的，在特定人群中相互之间也没有联系，更多的是随机出现。随机分配的特征，称之为连锁平衡（linkage quilibrium）。相反，表现为遗传连锁的遗传标记，比如那些在同一染色体上紧挨在一起的标记，称之为连锁不平衡性（linkage disequilibrium），在人群中，它们相互关联在一起的机率比预计的机率要高。

人体细胞含有 23 对（46 条）染色体，这 46 条染色体互相配对，配对的 2 个染色体被称为同源染色体，除性染色体 X 和 Y 外（Y 决定男性性别，比 X 小），所有配对（同源）的二个染色体大小、形状相同。在精子和卵子形成时，每个生殖细胞随机地接受配对染色体中的任何一套，即 23 条染色体，染色体之间不配对，为单倍体，当精子受精成为合子（受精卵），细胞拥有 46 条染色体，成为二倍体细胞。因此，每个人拥有二套染色体，一套来自父亲，一套来自母亲（图 2-4）。每个个体继承了父亲一半的遗传物质和母亲一半的遗传物质。在细胞分裂前，每个染色体一分为二，随着细胞的分裂，各自分配到子细胞中，由于染色体在细胞分裂过程中正确分配，每个子细胞接受相同的染色体，复制了父母的 46 条染色体，构成全部的遗传物质。因此，身体的各个体细胞含有相同的染色体组分，

这意味着各种不同组织，如血、毛发、皮肤、唾液等细胞含有相同的 DNA 信息，个体不同组织具有同一性，可以提供相同的遗传信息。

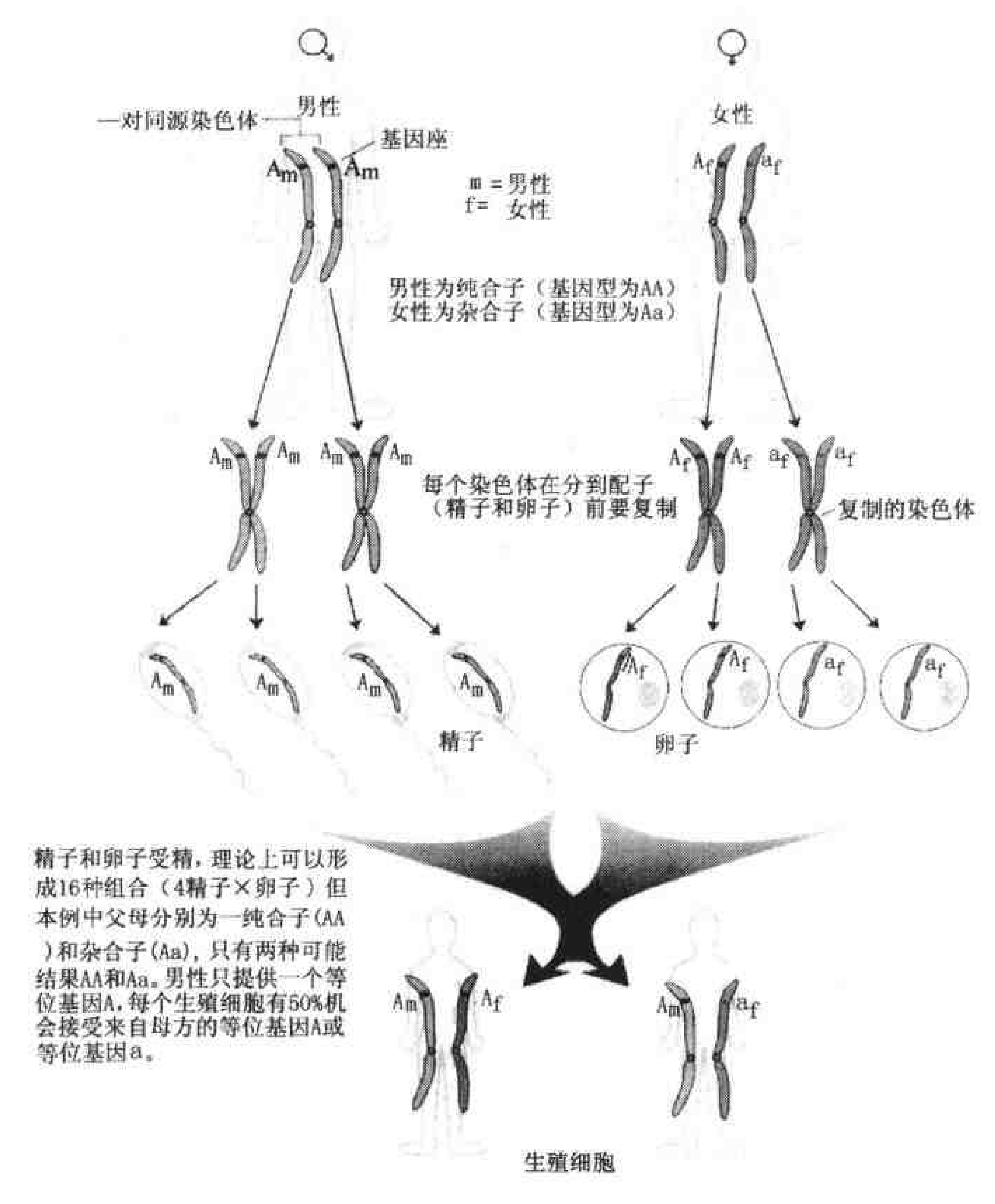


图 2－4 遗传图

精子细胞虽只含有 23 条染色体，即一半的染色体，但在分析中我们分析的是多个精子，包含有全部染色体。因此精液 DNA 能提供与血液或唾液等体细胞相同的遗传信息。

2.2.2 染色体结构

人类体细胞的 46 条染色体，其中 44 条（22 对）为常染色体，另两条与性别分化有关，为性染色体。女性性染色体为 XX，男性为 XY。生殖细胞均为 23 条染色体，其中卵细胞为 22 + X，而精细胞为 22 + X 或 22 + Y。人染色体根据他们的大小进行编号，1 号染色体最大，22 号染色体最小。

染色体在细胞周期中持续经历着凝缩和舒展的周期性变化。每个分裂中期的染色体都由两条染色单体所组成，各含一个DNA双螺旋分子，两条染色单体仅在着丝点（粒）互相连接，着丝粒是细胞在有丝分裂和减数分裂时染色体与由微管蛋白组成的纺锤丝相连接的部位，与细胞分裂时染色体的运动密切相关。染色体的基本区域如图2-5所示。染色体的中间区为着丝粒（centromere），它将染色体分成两部分，着丝粒的两端叫臂（arm）。臂的终端叫端粒（telomere）。短臂用“p”表示，长臂以“q”表示。一个完整的染色体至少包括着丝粒、染色体臂和端粒。

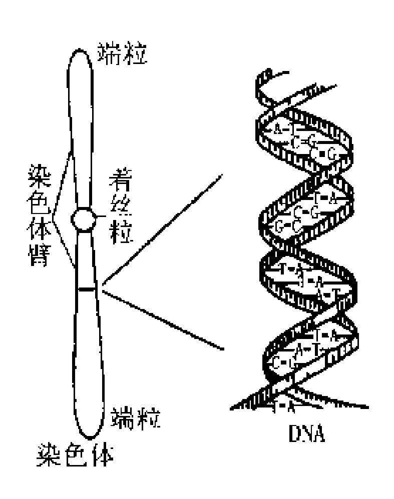


图2-5 染色体的基本结构

染色体臂是染色体的主体，它所包含的染色质可分为异染色质和常染色质两种：异染色质是包装紧密的局部区域，在细胞间期呈较大深染团块，异染色质主要分布于着丝粒末端和某些中间部分，一般不含基因或者很少有基因存在，所含DNA主要是一些高度重复序列。异染色质在不同的染色体上分布变化很大。

常染色质与异染色质不同，它占染色体的绝大部分，其表现具有染色质螺旋的正常周期，即在间期解螺旋，分裂时形成螺旋，在中期时达到螺旋高峰。但两者在化学组成上无差异，并且结构上相连，只是后者折叠较疏松。由于异染色质在遗传上是不活跃的，故染色体如果在这些区段发生断裂重接，甚至缺失或重复一段，对遗传影响并不大。基因主要位于常染色质区，故常染色质区的任何变化对遗传有重要影响。

端粒是真核生物线型染色体的末端，由DNA和蛋白质组成的复合体。它由一个简单重复序列串联排列而成，多为TTAGGG重复，紧接于简单重复序列的内侧，由中度重复序列组成的亚端粒区DNA称端粒相关DNA。端粒DNA长度在不同生物中变化很大，小鼠每条染色体含150kb端粒DNA，人类约10kb左右。端粒的复制需要一个特殊的逆转录酶即端粒酶。它是由一个RNA分子和一些蛋白质组成。复制时以其RNA中的一个片段为模板，以含端粒重复序列的寡核苷酸为引物，复制合成端粒序列。目前一般认为：DNA复制后，起始复制的RNA引物被降解，引物端留下的空缺由上游引物延伸合成的DNA序列填补替代。但与模板3′端配对的引物降解后，则没有引物延伸填补。因此，DNA每复制一次，其子链即缩短一次，当染色体端粒缩短至临界长度时，细胞即进人“程序化死亡”。端粒的功能是防止非活性的DNA末端降解和保证染色体的稳定性，同时使染色体末端之间不发生相互作用或融合（不加长或缩短）。随着年龄的增长或细胞分裂数目的增加，端粒的长度渐趋缩短。端粒酶在绝大多数机体组织、良性肿瘤和非永生细胞系中是无活性的，只在正常的生殖细胞、造血干细胞和恶性肿瘤细胞中有活性，因而认为端粒可能在恶性肿瘤的发生和维持中起重要作用。

2.2.3 染色体核型与物理定位

在多数细胞生活周期中，染色体以分散的线状形式存在。在细胞分裂前、有丝分裂（mitosis）的中期（metaphase），染色体浓缩形成一个紧密形式，染色后可以在显微镜下观察到。染色体经过一定程序处理与不同染料染色，在显微镜下观察，染色体由不同宽窄和

亮度的带连续交替组成，这种带叫染色体带，染色体上 A－T 与 G－C 含量不同造成带的宽窄和亮度不同。染色体带型染色常用 Giemsa 染色（简称 G 染色方法），因而称此带为 G－带（G－bands）（图 2－6），这些带在染色体高速公路上起着“信号灯”的作用，帮助确定那些特殊的 DNA 序列或基因与其他 DNA 标记在染色体上的位置。染色体大小和带型不同，可使 24 个染色体（22 个常染色体、X 和 Y 性染色体）彼此区分开，这种分型叫核型（karotype）。根据中期染色体的带型可以确定 DNA 或遗传标记在染色体上的位置，以着丝点为起点，根据特定染色体的短臂或长臂的相对位置横向分为若干区带和亚带。如 3P14.1 代表 3 号染色体短臂的 1 区 4 带的第 1 亚带；又如，染色体定位 12P1 表示 12 号染色体短臂上的 1 号带。带的数字按染色体的着丝粒向端粒部分递增，如带 32 比带 20 接近

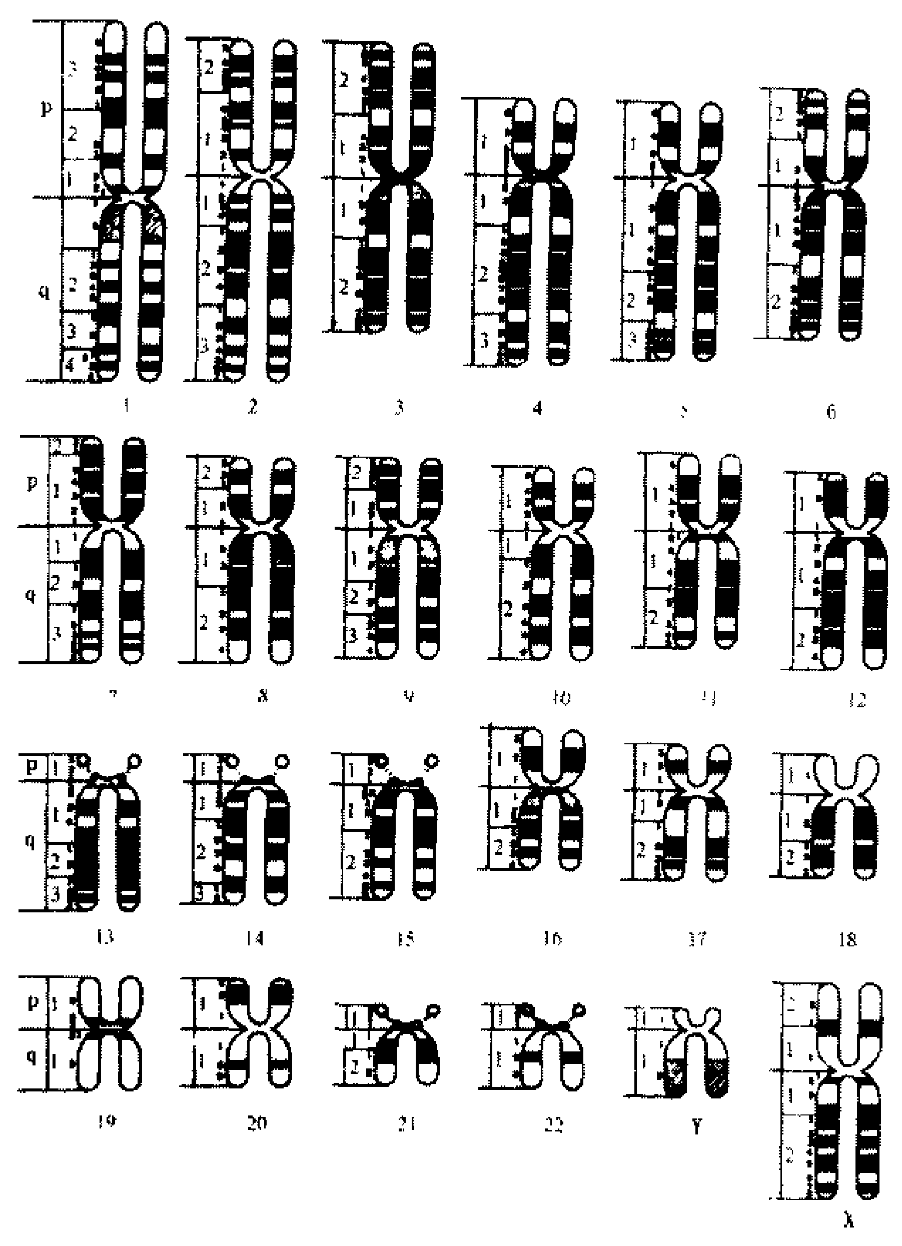


图 2－6　染色体核型与分带

端粒。位于染色体末端的 DNA 标记，命名时经常在染色体臂的命名后标上后缀“ter”，如 DNA 定位 15pter，表示位于 15 号染色体短臂的末端。有时 DNA 标记定位不是十分明确，在这种情况下，染色体定位表达为一个特定范围，如 2p23 - pter，意思是位于 2 号染色体短臂的 23 带与末端之间，从而为染色体的进一步识别、鉴定和基因定位创造了条件。

2.2.4 基因组

细胞或生物体的全套遗传物质叫基因组（genome）。所谓二倍体真核生物的基因组则是指维持配子或配子体正常功能的最基本的一套染色体及其所携带的全部基因。单倍体人类基因组由 3.2×10^9bp 的 DNA 组成，包括细胞核染色体基因组和细胞质内线粒体基因组（图 2 - 7）。细胞核染色体基因组为 23 条染色体 DNA，其中 90%左右 DNA 为程度不同的重复序列，10%为单拷贝基因。

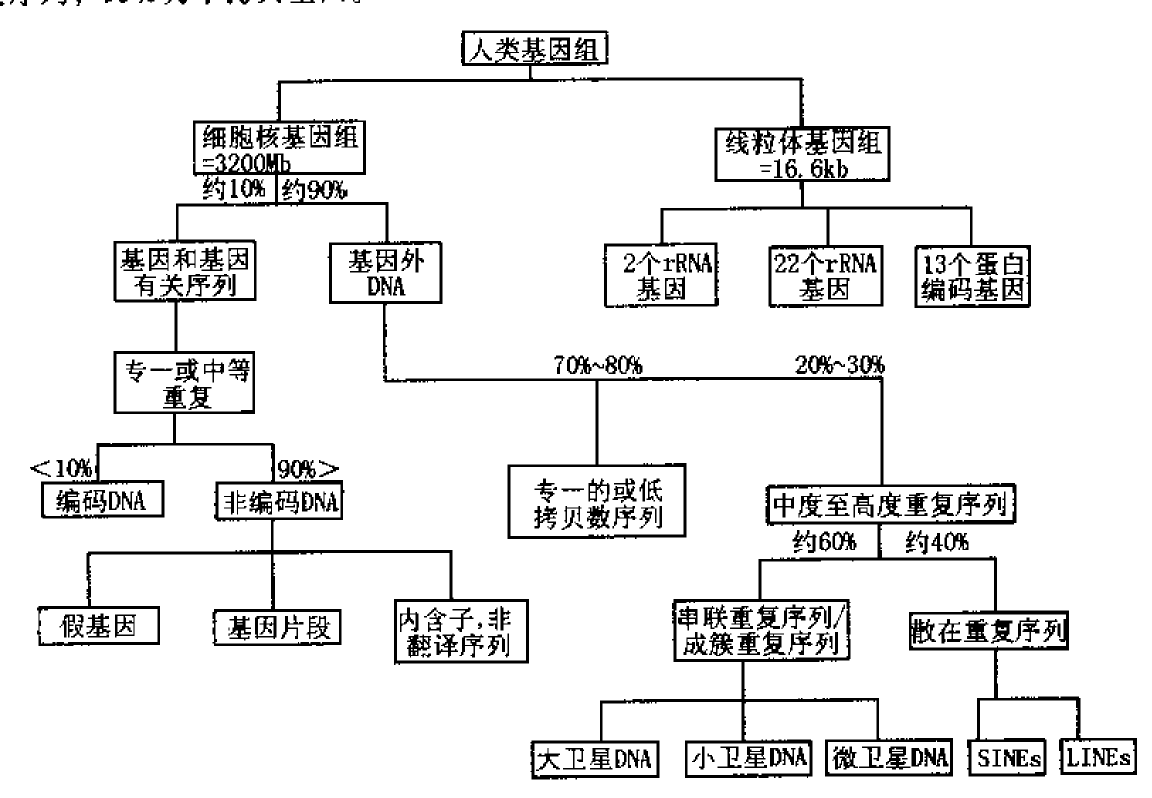


图 2 - 7　人类基因组 DNA

2.2.4.1 基因

染色体上的 DNA 物质由编码和非编码区组成，编码区域被称作基因，一个基因的大小一般为几个 kb 到几十个 kb。基因的概念随着遗传学、分子生物学、生物化学等领域的发展而不断完善。从分子生物学角度看，基因是负载特定生物遗传信息的 DNA 分子片段，在一定条件下，能够表达这种遗传信息，产生特定的生理功能。它包括合成一个有功能的多肽或 RNA 分子所必需的整个核苷酸序列，即除了蛋白质或 RNA 的编码区外，还包括为获得一个特异转录产物所必需的其他 DNA 序列，如一些非编码序列。

1. 基因的结构

一个基因包括 3 个区域：

(1) 编码区：包括编码区（外显子）和内含子；

(2) 前导区：位于编码区上游，相当于 mRNA 5′端非编码区（非翻译区）；

(3) 调节区：包括启动子和增强子等基因编码区的两侧区，也称侧翼序列。

2. 基因的分类：

按功能基因可分为结构基因和调控基因。

(1) 结构基因：可被转录形成 mRNA，并进而转译成多肽链，构成各种结构蛋白质、催化各种生化反应的酶和激素等；

(2) 调控基因：指某些可调节、控制结构基因表达的基因。其突变可影响一个或多个结构基因的功能，或导致一个或多个蛋白质（或酶）量的改变。此外，还有一些只转录而不翻译的基因，如核糖体 RNA 基因，tRNA 基因。

2.2.4.2 重复序列

人类基因组超过 90%的序列为重复序列，它们不编码 mRNA 前体或其他 RNA。个体之间重复序列的结构相似但并不相同，而表现出巨大差异性，依据这些差异通过 DNA 指纹等技术可进行区分。

根据重复序列的重复次数可将其分为 3 类：

1. 高度重复序列：这类 DNA 复性极快，由数种不太长的核苷酸序列重复排列，长度从几个 bp 到几百 bp 或更长，重复次数 $>10^6$，占总含量 1% ~ 50%。它们具有下列特点：

(1) 其中的一部分常以“祖先或原始序列片段”形式出现，进化上是原始的，趋向保守，重复序列的另一部分则以新进化的家族形式出现。

(2) 不同家族常源于同一个“祖先重复序列”，结构有其相似性，但同一家族中各个体间又有不断积累的个体突变。

(3) 凡高度重复序列多呈线状顺向排列，以一定的有序组织排列在基因组中，有各自特有的重复周期，甚至限制酶切点也有相应分布。同一原始来源的各重复序列拥有相似重复周期。

(4) 新的高重复序列家族，往往是在新的物种衍化或在进化之中，通过一定机制（如基因的不等位交换）形成，随基因的扩大而被放大。

基因组中由寡核苷酸串联、重复排列而成的 DNA 序列，构成数量可变的串联重复序列。由于该区 AT 含量高，如果将基因组 DNA 切成片段后，在氯化铯（CsCl）溶液中进行密度梯度离心时，可在一个主区带以外形成小的区带，称为卫星 DNA。其中重复次数为 6 ~ 70bp 的称为小卫星，它常位于染色体近端粒和着丝粒区，常带有 GGTGGGCAGARG（R：嘌呤）的核心同源序列。而在基因组的间隔序列和内含子等非编码区内，广泛存在着与小卫星 DNA 相似的一类小重复单位，重复序列为 2 ~ 6bp，称为微卫星 DNA，这些微卫星 DNA 在人类基因组中出现的数目和频率不同，表现出多态性。

2. 中等重复序列（intermediate - repeat DNA）：亦称转座元件（transposable element）。它是一种能从基因组的一个位置移动到另一无关位置的 DNA 片段，插入后可影响所在位置基因的表达。转座元件的长度一般为 2 000 ~ 12 000bp，其拷贝数一般在 10 ~ 10^5 之间，占整个基因组的 25% ~ 40%，复性速度要慢一些，一般是非编码序列。在人基因组已知有两大家族，即长散在成分或叫长散在重复（long interspersed nucleotide element，LINEs）和短

散在成分或叫短散在重复（short interspersed nucleotide element，SINEs）。

3. 低重复系列：拷贝数为 2～100，复性速度很慢。

2.2.4.3 AluⅠ序列

AluⅠ序列为短的散在重复序列，该序列中有一个核酸内切酶 AluⅠ（识别序列为 AC-CT）的识别位置。广泛存在于哺乳类动物中，占人类基因组总 DNA 量 5%左右，高达 9×10^5 拷贝，主要集中在细胞分裂晚期的 R 带，大部分属非编码 DNA，但也有一部分位于 mRNA 的非翻译区，甚至位于编码区内。

人类 Alu 序列长度约 300bp，本身又由 120bp 和 150bp 的重复序列组成，两者之间由富含 A 的区域分开，两端又有一段 7～10bp 的正向重复序列。它可能是 7sRNA 降解形成并反转录后整合于基因组中，从而在体外对许多启动子发挥正性或负性转录调节作用，但这些序列在体内条件下可能由于 DNA 甲基化和/或处于核糖体部位而不能转录。Alu 序列在体细胞中几乎完全甲基化而在精子中处于低甲基化状态。

人类基因组内至少存在 4 种不同类型的 AluⅠ序列，属于两个不同的亚家族。一些类型的 Alu 是人类特有的，并表现为二态性（dimophic）。Milewilz DM 证明人类 3 型胶原纤维基因的第 8 个内含子中存在一类种族特异的 AluⅠ序列，35%的美国黑人中存在这类序列，1%的高家索人有这类序列，而印第安人、东南亚人则不存在这类序列。黑猩猩、绿猴、长臂猿等非人灵长类中均无此序列。在 Y 染色体中有一人类特异的 AluⅠ序列，称为 YAP（Y Alu polymorphism），在不同人群中表现出显著的差异性。

2.2.4.4 长散在重复成分

长散在重复成分在人类基因组中约有 10^5 拷贝，占总 DNA 的 2%～3%，长度为 6～7kb，其中 95%的序列 5′端是截断的，但大多数含有相同的 3′端及长短不等的 PolyA。可能由于内部缺失或重排，每个长散在重复成分与同源序列约有 13%的差异。长散在重复成分均位于常染色体的 G/Q 带，其中可表达的基因极少，可能代表由逆转录形成并重新整合进人染色体的假基因样拷贝。

多数人认为重复序列是有功能的。如：可能参与维持着丝点的强度，形成结构基因间隔；参与减数分裂时同源染色体的联会配对；参与某些基因的重组，保证染色体的三维结构；使基因组容易重组及造成新物种的形成。总之，对于重复序列的功能仍有待进一步研究。

2.2.4.5 等位基因

相同基因或标记的不同形式叫等位基因（allele），基因或非编码区的 DNA 标记在染色体上的位置叫基因座（locus，复数形式为 loci），当等位基因的频率大于 0.1 时，就称之为有多态性。多态性（可变）标记广泛分布在人基因组的非编码区。等位基因最简单的例子是 1866 年由现代遗传学之父－乔治·孟德尔观察到的豌豆有绿色和黄色，圆粒和皱粒，孟德尔通过对豌豆数代这些特征的观察，发现绿色和黄色是同一基因的相互变化的形式（等位基因），圆粒和皱粒是另一基因的不同等位基因。ABO 血型是人类同一基因的不同等位基因。在基因组某一基因座的等位基因在一对染色体上是相同的话，这种情况就叫纯合的（homozygous），在 DNA 图谱上表现为一条带；相反的，如果在特定基因座表现有差异性，则称为杂合的（heterozygous），在 DNA 图谱上表现为二条带，这些差异可能或不可能在物

理外观上显现出来。基因型（genotype）是某个体在一特定基因座的等位基因的特征。当配子（Gametes）（卵子和精子）形成时，每一个配子随机地得到父亲或母亲的一半遗传成分（图2-4），染色体分配结果使每个配子正确含有每一个染色体的一份拷贝，总共23条染色体，这意味着如果父亲或母亲在某一基因座是杂合的话，只有二个可能等位基因的一个到每一个配子中，因此同一基因的等位基因彼此分离。这表示一个个体从父母那里会继承二个同源的染色体，如二个6号染色体，一个来自父亲，另一个来自母亲；同样地，二个7号染色体，也分别来自于父母双方，但每个个体不可能得到来自父亲的二个6号染色体或7号染色体，也不可能得到来自母亲的二个6号染色体或7号染色体，这个原则叫独立分离原则（independent segregation）。这是保证子代最终拥有正确的染色体数（2×23条）的机制。独立分配和随机分配原则促进了物种的不断分化和遗传的稳定性。

2.3 DNA 的二种多态性（差异性）

人们发现，碱基在生物进化过程中，由于碱基突变、缺失、插人或置换，DNA 分子在人群间有差异，这些发现是在疾病调查研究中得到的，单一基因座的遗传标记有多个等位基因存在，就叫多态性（polymorphism）。当这样的基因座表现出有很多的变异体（多达几百个）时，就称之为高度变异性（hypervariable）。

DNA 的多态性或差异性可分为二种：（1）表现在碱基序列上，即序列多态性；（2）表现在二个固定端点间 DNA 片段的长度上，即片段长度多态性。序列多态性就像英国英语和美国英语中同一个单词不同的拼写方法，当我们看到 analyze 拼写成 analyse，单词和意思我们仍能认得，单词词根也同样，比如二个双链 DNA 片段

AGCTCAATCA		AGATCAATCG
TCGAGTTAGC	和	TCTAGTTAGC

从左起第三个碱基处表现出序列多态性，其余碱基序列相同。

长度多态性（length polymorphisms）（图2-8），最容易比作载有不同车厢的火车，火车头和车尾是火车的二个固定点，整个长度随所接车厢的数目而变，下面是三个车厢或连续重复 DNA 序列的例子。

AGCTCAATCA - AGATCAATCG - AGCTCAATCG - R

TCGAGTTAGC - TCTAGTTAGC - TCGAGTTAGC - R

按遗传学术语，车厢叫串联重复，有串联重复数差异性的基因座称为可变数目串联重复（Variable Number Tandem Repeat，VNTR）基因座，等位基因含有的串联重复数目，可以作为该等位基因的命名，如串联重复数为35，就命名为“等位基因35”。

2.4 遗传标记分类

基因组 DNA 有两种差异性，由此产生的遗传标记也分两大类：长度多态性标记和序列多态性标记。前者根据分析检测的方法又分为限制性内切酶片段长度多态性和扩增片段长度多态性。此外，有一类是既有串联重复单位数目变化，重复单位内部又夹杂有序列的变化，即既具有长度多态性，又有序列多态性，如 D1S8 基因座，称此类为小卫星可变重复单位多态性。因此，从这个角度，遗传标记共有三大类（图2-9）。

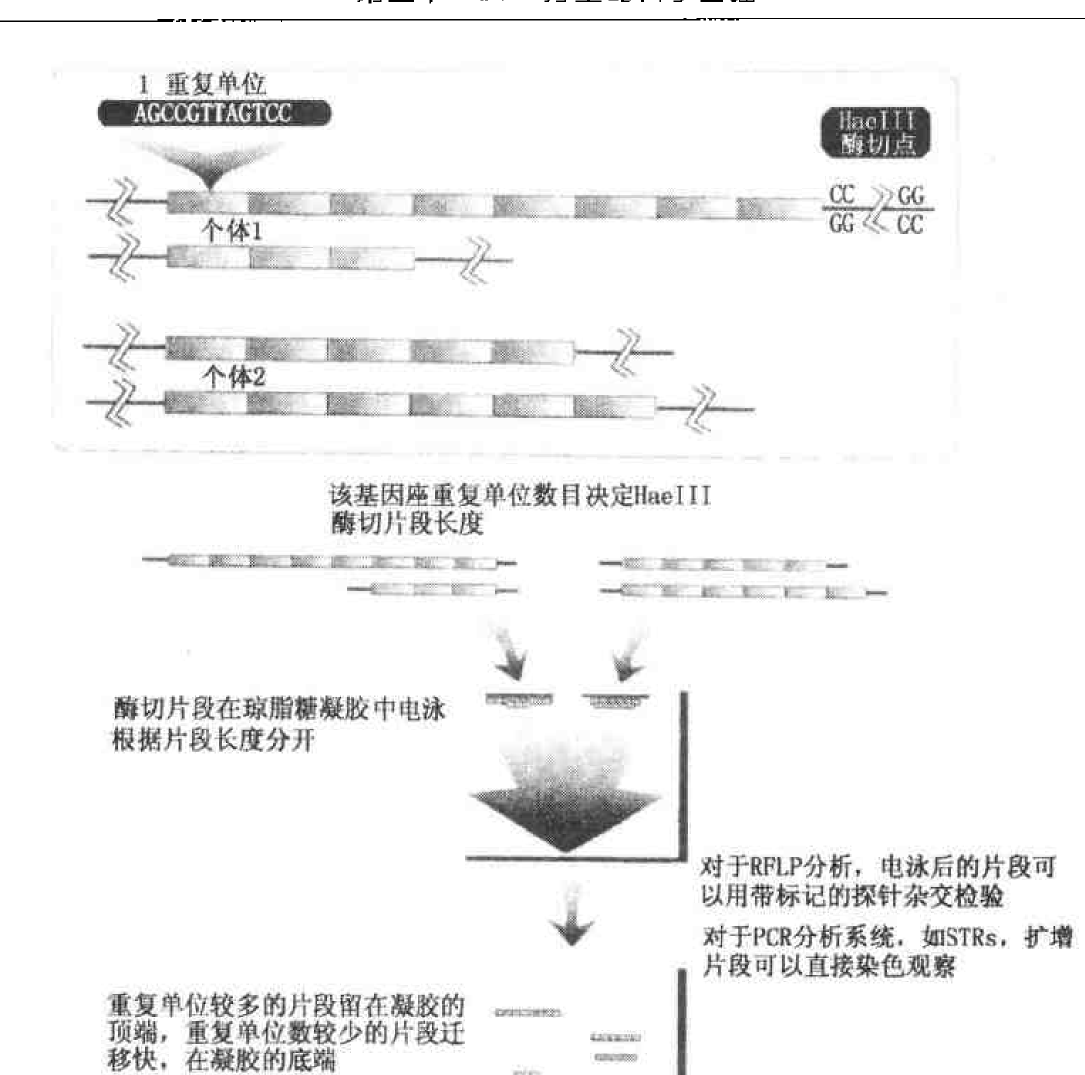


图 2－8　长度多态性

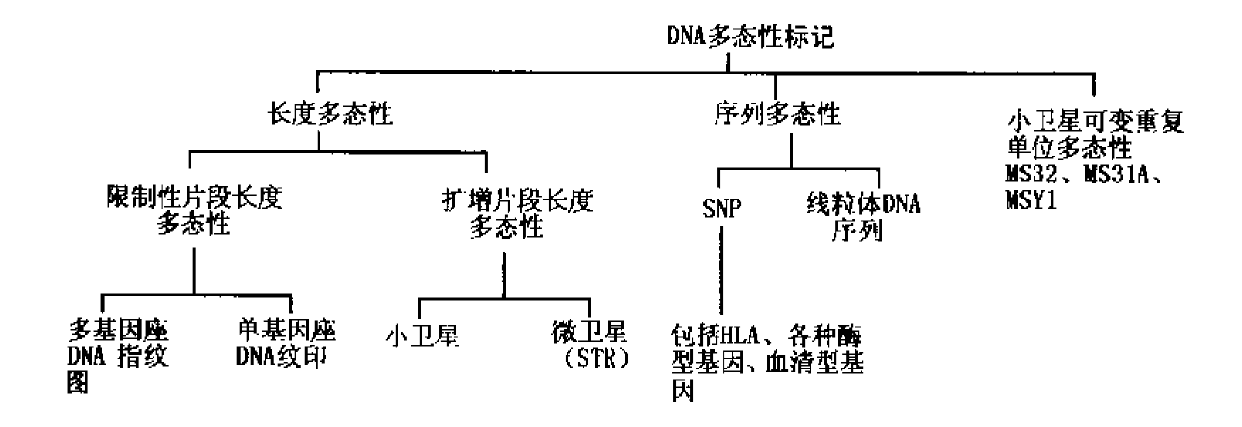


图 2－9　遗传标记分类

2.4.1 小卫星与微卫星

目前常用的长度多态性标记是小卫星和微卫星，其中重复单位长度为 6 ~ 70bp 的串联重复序列，称为小卫星（minisatellite）DNA，1980 年 Wyman 和 White 在人类基因组中发现了第一个这样的小卫星序列。串联重复单位的长度为 2 ~ 6bp 的重复序列被称为微卫星（microsatellite）DNA，这是 Weber 在 1989 年通过 PCR 扩增和直接测序发现的。由于小卫星和微卫星 DNA 在人类基因组中出现的数目和频率不同，表现出多态性，因而为人类遗传分析提供了大量的多态遗传标记。小卫星和微卫星 DNA 各有其特点，表 2 - 2 为两者比较。小卫星和微卫星 DNA 的多态性产生机制不同，小卫星的多态性是有丝分裂、减数分期姐妹染色单体不等交换或染色体内部不等交换的结果。微卫星的多态性主要是 DNA 复制过程中滑动，或 DNA 复制和修复时滑动链与互补链碱基错配，导致一个或几个重复单位的缺失或插入的结果。微卫星 DNA 是目前最常用的遗传标记。

表 2 - 2 小卫星 DNA 和微卫星 DNA 的区别

	微卫星 DNA	小卫星 DNA
存在部位	染色体任何部位	染色体近端粒和着丝粒区
重复单位长度	1 ~ 6bp	6 ~ 7bp，常富含 GC
重复次数	10 ~ 60 次	几次到几百次
总序列长度	约 200bp	0.5 ~ 30kb
重复单位的差异	重复单位的变异性低，可看成结构相同	单位组成稍有差异，如单个碱基置换
存在数量	很多，整个人基因组 5 ~ 10 万	有限，有些染色体尚未见到

2.4.2 单核苷酸多态性

序列多态性是指在两条同源染色体上，同源 DNA 序列长度相等，但个别核苷酸存在差别。由单个核苷酸替代、插入或缺失而形成的分子多态，又称为单核苷酸多态性（single nucleotide polymorphisms，SNPs）。

单核苷酸多态性表现为一种双或二等位基因形式（在基因组中某基因座要么突变，要么正常），即特定核苷酸位置上存在 2 种碱基形成。原因大多数是替换，且多为转换，即一种嘌呤碱替换另一种嘌呤，或一种嘧啶碱替换另一嘧啶如 A 和 G 之间或 T 和 C 之间的替换。少数为颠换，即嘌呤与嘧啶的互换。在单核苷酸多态基因座，转换与颠换的比例大约是 2:1。

进入 20 世纪 90 年代，随着人类基因组计划不断深入，单核苷酸多态性才逐渐引起重视。SNPs 是人类基因组中数量最多的一种多态形式，人类基因组中共有 300 万个 SNP，平均每 1 000 个碱基中就有一个 SNP，比 STR 多出几个数量级。单核苷酸变异最常见，分布最广。单核苷酸多态性最能反映出群体遗传结构特征。单核苷酸这类遗传标记更有助于解释个体间差异，以及不同群体、不同个体对疾病的易感性。SNPs 被称为第三代遗传标记。

DNA 芯片技术的建立和应用，使在不久的将来可以用足够量的 SNPs 构建第三代高密度的人类基因图谱，阐明人类表型差异以及对疾病易感性的分子机理。

2.4.3 DNA 标记（marker）的命名

DNA 标记的命名，国际上有统一的标准。具体是：

对于是基因的一部分或位于基因内部的标记，用该基因的名字命名。例如，STR 标记

TH01 是位于 11 号染色体上的酪氨酸羟化酶基因，TH01 中的“01”意味着所涉及的重复区定位于酪氨酸羟化酶基因第 1 个内含子。有时在基因座名字前加一个前缀“HUM”，这表示来自于人基因组，因此 STR 基因座 TH01 正确地表示为 HUMTH01。

对于基因区外的 DNA 标记，以他们所在的染色体位置命名。例如，D5S818 和 DYS19。他们不在编码区。这里，“D”代表 DNA，D 后的数字代表染色体的编号，5 指 5 号染色体，Y 指 Y 染色体。“S”指的是 DNA 标记为单（single）拷贝序列，最后的数字代表该标记发现的顺序和特定染色体的分类。

用有序数目来命名每个 DNA 标记，使每个 DNA 标记独一无二，不会重复。

例如，D16S539 标记的每个符号表示：

D：DNA

16：染色体 16

S：单拷贝序列

539：染色体 16 上第 539 个基因座

2.5 多态性产生的机理

2.5.1 多态性产生过程

DNA 多态性产生于以下 4 种生物学过程：（1）细胞分裂间期，DNA 复制过程；（2）减数分裂前期，同源染色体联会后，非姐妹染色单体间的交叉，重组与互换过程；（3）减数分裂中期非同源染色体的自由组合；（4）受精过程中配子间的随机结合。

基因组结构的相对稳定是生物种系得以维持和延续的基本前题和保证，细胞内在的一整套的修复系统以保证遗传忠实性。但在生物进化过程中，会出现偶然的 DNA 错误如单个核苷酸的替换（substitution）、单个或多个核苷酸的缺失（deletion）和插入（insertion），两个或多个碱基对的双链 DNA 片段转动 180°即“倒位”（inversion）和一段 DNA 序列的双倍重复（duplication），所有这些导致生物体个体之间有差异，表现出多态性。核苷酸的替换有转换（transition）和颠换（transversion）两类。这种单个核苷酸替换使密码子发生改变形成的单核苷酸多态，在基因组中最为常见。

2.5.2 多态性形成原因

1. 不等位交换

人类基因组中，碱基缺失和插入是一种较为常见的突变形式。缺失和插入可能由几种机制造成，机制之一是不等位交换（unequal crossing over）。图 2－10 为一个简单的模型，按此方式两条染色体间不等位交换，结果在一条染色体上有一片段缺失，而在另一条上则出现相应的添加。一般来说，较大片段的插入和缺失是由这种机制造成的。

2. 复制滑脱

复制滑脱（replication slippage）或链滑脱误配（slipped－strand mispairing）是造成缺失和插入的第二种机制。这类发生在邻接串联重复序列的 DNA 区域中，图 2－11 所示，而滑脱又可造成某一片段 DNA 缺失或重复，至于究竟是缺失还是重复则要看滑脱发生在 5′→3′方向还是与之相反的方向。在 DNA 复制过程，复制滑脱可能导致一些疾病。在癌细胞中，微卫星 DNA 的串联重复数也明显增加，有些癌细胞中，串联重复数可增加 100 倍，

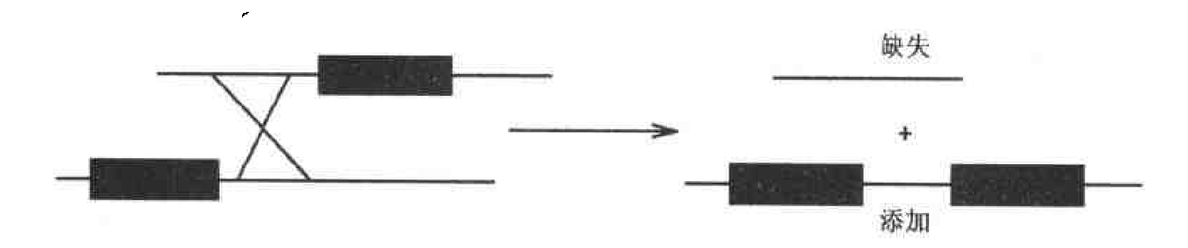


图 2-10 不等位交换

两条子链某一 DNA 序列缺失，而另一子链则出现同一序列的重复，图中矩形表示某一特定长度的 DNA

不仅 DNA 链在复制时滑动，DNA 聚合酶也会滑回到原先已复制过的 DNA 区段，串联重复数迅速增加。一般认为，微卫星串联重复主要由复制滑脱产生的，而较长片段的串联重复如小卫星 DNA 则主要由染色体不等价交换造成。在统计学上，微卫星串联重复可采用同步突变模式（stepwise mutation model，SMM），即每一次突变都增加一个或数个重复序列，产生一个新的等位基因。按此模型计算微卫星的突变率为 10^{-4} ~ 10^{-5}，且核心序列碱基数越小，突变率越低。小卫星串联重复可采用无限突变模式（infinite mutation model，IMM），即每一次突变都产生了一个在以往群体中未被发现的新的等位基因。与此模式相符，小卫星 DNA 的突变率为 5×10^{-2}。

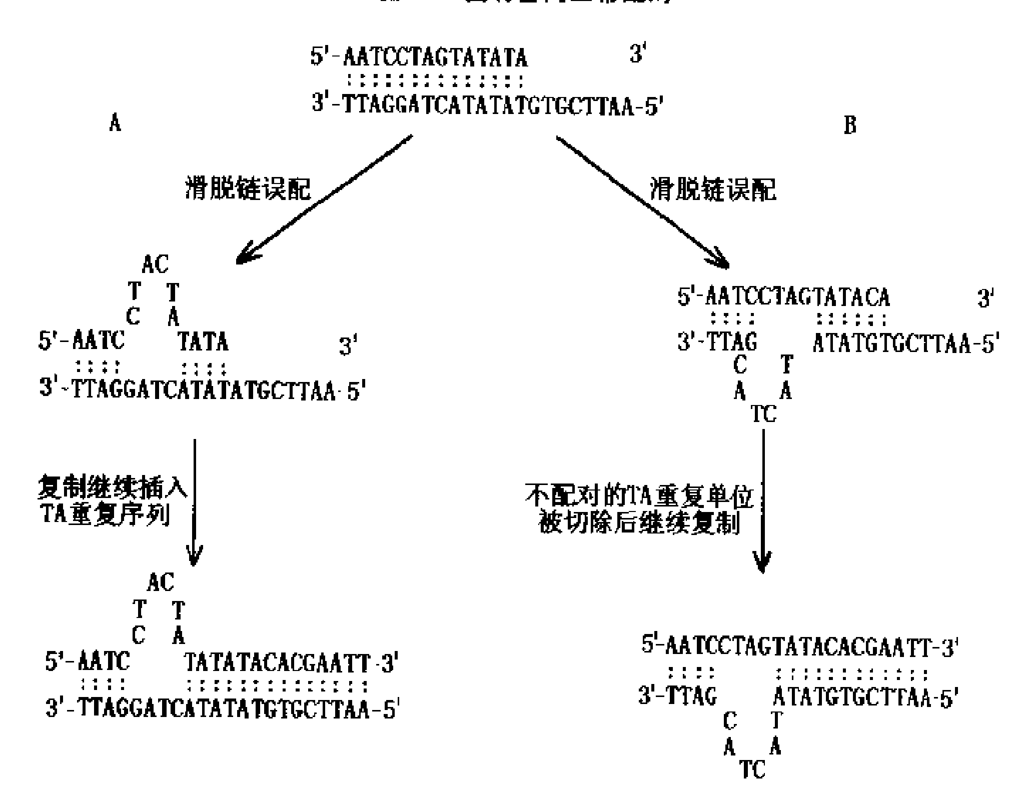


图 2-11 滑脱误配所产生的重复和缺失

小箭头表示 DNA 合成的方向，圆点表示碱基配对；(A) 在 DNA 复制期间 TA 重复中发生的两碱基滑脱：3'→5'方向中的滑脱造成一个 TA 单位插入的结果；(B) 另一个方向中的滑脱则造成一个重复单位缺失的结果，缺失正在延长的链的 3'端上，未配对重复单位是被切除而造成的，可能是凭借了 DNA 合成酶的 3'→5'外切酶活性。

3. 转座和反转录

转座（transposition）和反转录转座（retroposition）是另外一种导致突变的形式。遗传物质从一个染色体位置向另一个位置的运动叫转座。以 RNA 为中介的转座形式叫反转录转座。

有丝分裂前期，每条染色体含有两条姐妹染色单体，同源染色体配对后，非姐妹染色单体发生互换，结果等位基因间组合形式发生改变，如图 2－12 所示。A/a、B/b 是两个座位的等位基因，A 和 b 在一条染色体上，a 和 B 在另一条染色体上。两个基因座的等位基因是否在互换中发生重新组合，与两个基因座的连锁距离有关，两个基因座相距越近，重组的机率越小。染色体互换、重组导致配子间的基因型不同，从而出现后代的表型差异，即使同胞姐妹，在表型上也会出现差异。

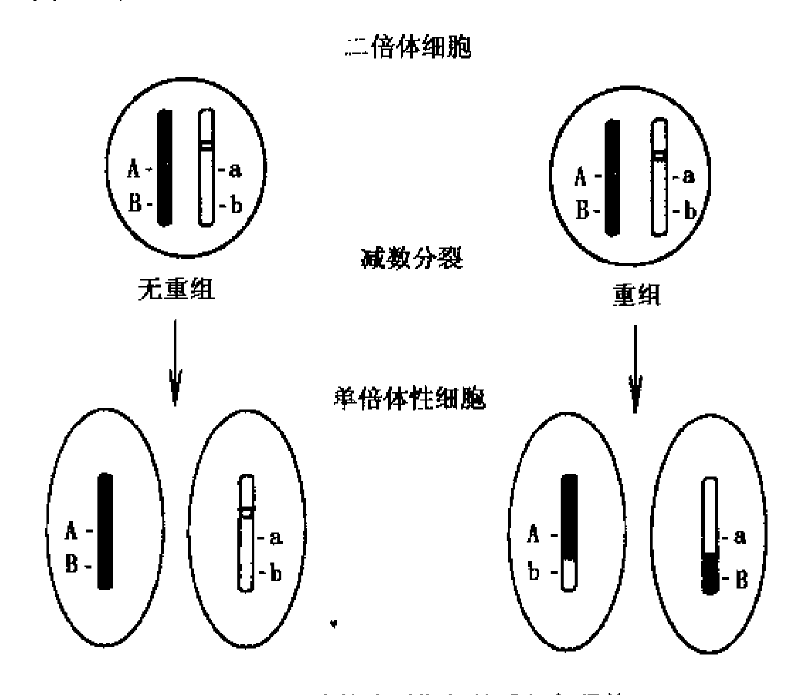


图 2－12　减数分裂期间的重组与互换

有丝分裂后期，同源染色体对分离，不同的同源染色体在分向两极时，相互间是独立的，因而父方、母方来源的染色体有机组合，产生基因组变异。人有 23 对染色体，父母双方之间有 2^{23} 的组合形式，由于还有父母染色单体的重组混合，因而除了同卵双生(identical twins)外，几乎不可能得到遗传上相同的后代。另外，在受精过程中，具有不同基因组合形式的无数个精子中，只有一个与卵子结合，这种结合也是随机的，理论上，任何一种基因组合形式的精子都有平等的机会与卵子结合，因此这一过程也可产生基因多样性。

2.6 群体遗传

等位基因频率和基因型频率是群体遗传的遗传结构。等位基因频率（allele frequency）指在一个二倍体的某特定基因座上某一个等位基因占该基因座的等位基因总数的比率。这

是群体遗传结构的一个最基本测度。基因型频率（genotype frequency），指的是某特定基因型个体的数目占总数目的比率。

所有遗传标记在不同种族人群中频率是不同的。众所周知，ABO 血型的频率不同，比如在高加索人群中，B 型血型大约占 10%。DNA 分析中所用到的遗传标记的每一个等位基因也表现出不同人群的频率差异，因此确定遗传标记的出现频率对了解某一基因型的意义是很重要的。例如，某一基因型在 50% 的人群中出现，如果犯罪现场样品和嫌疑人都拥有这一基因型，就显得不那么特殊，没有多大意义，因为本来就有 50% 的人就有可能相同，也就是说任意抽出 100 人中，就有 50 个人彼此相同，个人特异性就很低；反之，如果这基因型在人群中出现频率为百万分之一，那么现场样品和嫌疑人都为这一基因型，其意义就很大，此嫌疑人就是现场样品遗留者的可能性或准确性就很大。

在生物物证鉴定中，一般要分析检验多个遗传标记，遗传一致性的总体意义的最有效表达方法是各个标记不同频率的乘积，每增加一个标记，如果相同，遗传一致性就会随着提高。但是所用的遗传标记必须预先通过遗传和统计学分析，这样的计算方法才是合理的。这要求所用的遗传标记研究的群体首先符合哈德 - 温勃格平衡（Hardy - Weinberg equilibrium，简称 HW），表示某一基因座等位基因与另一等位基因没有优先的相关性，它是群体遗传学的基石。如有相关性，就会人为地夸大遗传一致性的意义。HW 平衡要求在足够大的随机人群中，没有因迁移、自然选择或基因突变的发生，下一代基因型的频率和上一代一样，没有大的变化。第二点是，不同位点的基因分型之间没有相互联系（或相关性），这就是著名的连锁平衡，只有当群体有很多的亚结构（Substructure）或亚种族时，这个人群才会偏离 HW 平衡或和连锁平衡，如果在一个较大的混合群体中少数部分彼此间婚配，产生一个相对独立的生殖群时就会发生上述现象，比如，云南少数民族中，民族间表达出极高的内部联姻，这样群体中得到的等位基因频率就会与从大的混合群体中得到的频率偏离，因此亚种族的任二个个体拥有相同遗传标记的概率与用普通群体调查中得到的等位基因频率预测的概率不同。如果所涉及到的等位基因在亚群体中频率很高，遗传图谱相似性意义就会过高估计。在任何条件下，必须进行群体的样品调查以获得等位基因频率。

哈德 - 温勃格平衡是在 1908 年由英国数学家 G.H.Hardy 和德国医生 W.Weinberg 分别独立发现的，故称为 Hardy - Weinberg quilbrium，有时也叫定律。HW 定律的 3 个基本点：

第一，等位基因的频率世代相传不会改变。常染色体上的一对等位基因 A 和 a 的频率分别为 p 和 q，且 $p + q = 1$。

在群体中这一对基因的 3 种可能的基因型频率为：$D = p^2$，$H = 2pq$，$R = q^2$。在下一代中，基因型的频率将与前一代的频率相同，则这个群体（p^2，$2pq$，q^2）称为平衡群体，或以 $(pA + qa)\ 2 = p^2\ (AA) + 2pq\ (Aa) + q^2\ (aa)$ 二项式展开式表示。因此，HW 定律有时称为二项式平方定律（biomial square law）。

以表 2 - 3 中的基因和基因型作为亲代群体的基因和基因型频率，下一代群体中的等位基因（A）频率为：$p1 = p^2 + 1/2 \times 2pq = p\ (p + q) = p$；同理等位基因 a 的频率为：$q1 = q^2 + 1/2 \times 2pq = q\ (p + q) = q$

第二，二项式展开式得到平衡的基因型频率，而且也不改变。因为下一代群体中的基

因频率仍为 p 和 q，如同在亲代群体中一样，3 种可能的基因型的频率仍分别为 $D1 = p^2$ (AA)、$H1 = 2pq$ (Aa)，$R1 = q^2$ (aa)。下一代基因型频率与前一代一样，基因型的频率平衡在 (p^2, 2pq, q^2) 状态之中。

第三，只需通过一个世代的随机交配就能达到基因型频率的平衡。

表 2-3　等位基因 (A, a) 的 Hardy-Weinberg 平衡

雌配子及其频率	雄配子及其频率	
	A (p)	a (q)
A (p)	AA p^2	Aa pq
a (q)	Aa pq	aa q^2

2.7 法医 DNA 分析中常用的两种酶

不论在体内或体外，如果不提到酶，不可能讨论生化反应，酶是使工作做完的生化工作者，酶以蛋白质为基础、能够多次反复催化形成或裂解其他生物物质。在实验室中，从细胞裂解到 DNA 合成的一切过程都要用到各种酶。如果没有酶，也无现代分子生物学（或任何生物学）可言。酶在体内温度和合适的条件下发挥作用，一旦从自然条件下分离出来，要求在试管中发挥酶的催化作用，创造与他们自然环境尽可能接近的条件是十分关键的，否则的话，酶不会发挥预期的作用。

在 DNA 分析中涉及许多酶，最常用的有二类酶，一类是多聚酶，另一类是限制酶。

多聚酶，它在 DNA 复制中引导核苷酸的增加，就是这种酶使 PCR 成为可能，Taq DNA 多聚酶拥有即使在加热条件下仍保持其活力特异的特点，这点对 PCR 十分重要。

限制酶，它能识别特定的 DNA 序列，并在其切点将 DNA 切成较小的片段。限制性酶由细菌体内产生，以保护细菌免受病毒侵入。它通过识别病毒中存在的特定的小 DNA 序列，在所有存在这种序列的地方切断病毒 DNA，而细菌本身 DNA 通过机体其他生化机制先切成小片段而保护起来。

第三章 物证的搜集和保存

成功的 DNA 分析始于生物物证检材的正确提取、保存与送检。送入实验室检验的物证 DNA 的状况，决定着 DNA 分析方法的选择、有效结果的获得以及结果解释与应用。了解 DNA 在各种生物检材中的含量、DNA 离体后的变化、各种外界因素对 DNA 的影响、现有分析方法对 DNA 检材的要求以及 DNA 的变化和外界因素对检验结果的影响，有利于正确提取、保存和送检生物检材。

3.1 生物物证中 DNA 的含量

人体细胞内有两种 DNA：一是核基因组 DNA，存在于细胞核内，二是线粒体基因组 DNA，存在于细胞器内。目前法医 DNA 技术可以分析这两种基因组 DNA 的多态性，理论上，有细胞的生物物证均可进行 DNA 分析，常见生物检材如下：

（1）血液和血斑；

（2）精液和精斑；

（3）各种组织和器官；

（4）骨骼和牙齿；

（5）毛发和指甲；

（6）唾液、尿液和其他生物体液（斑）。

不同的生物样品，DNA 含量不同，表 3-1 列出了一些常见的生物样品可以提取出的核 DNA 数量。精液中 DNA 含量很高，而脱落的毛根 DNA 含量很少。各种组织 DNA 含量顺序是肾 > 淋巴结 > 脾 > 肝 > 脑 > 肌肉 > 心 > 血（图 3-1），每克肾组织 DNA 可达 800 微克。但各种组织内的 DNA 降解速度不一，有些组织 DNA 易降解，保存一段时间后组织 DNA 含量发生较大变化（图 3-2，3）。脑组织的 DNA 相对稳定，在低温无菌条件下保存一个月 DNA 仍没有明显减少，保存一个月以上，DNA 也只减少一半，仍可达 40μg/g（图 3-4）。脑组织 DNA 保持较好的主要原因是脑存在于颅腔内，不易受到环境影响和细菌的侵害。

表 3-1 生物样品中的 DNA 含量

样品类型	DNA 含量
血液	20 000 ~ 40 000ng/ml
血斑（1 cm^2）	250 ~ 500ng
精液	150 000 ~ 300 000ng/ml
性交后阴道拭子	10 ~ 3 000ng
拔下的毛发	1 ~ 750ng/根
脱落的毛发	1 ~ 12ng/根
唾液	1 000 ~ 10 000ng/ml
口腔拭子	1 000 ~ 1 500ng
尿	1 ~ 20ng/ml
骨骼	3 ~ 10ng/mg

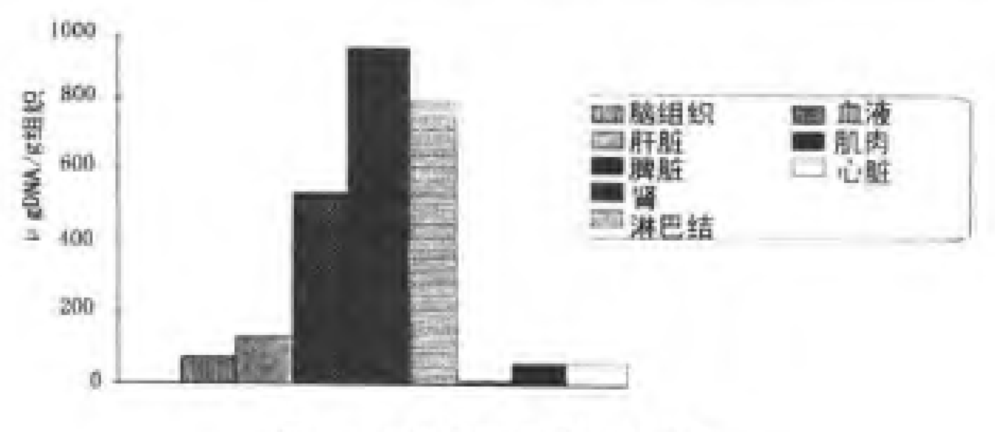


图 3-1　组织采集一天后的 DNA 产量

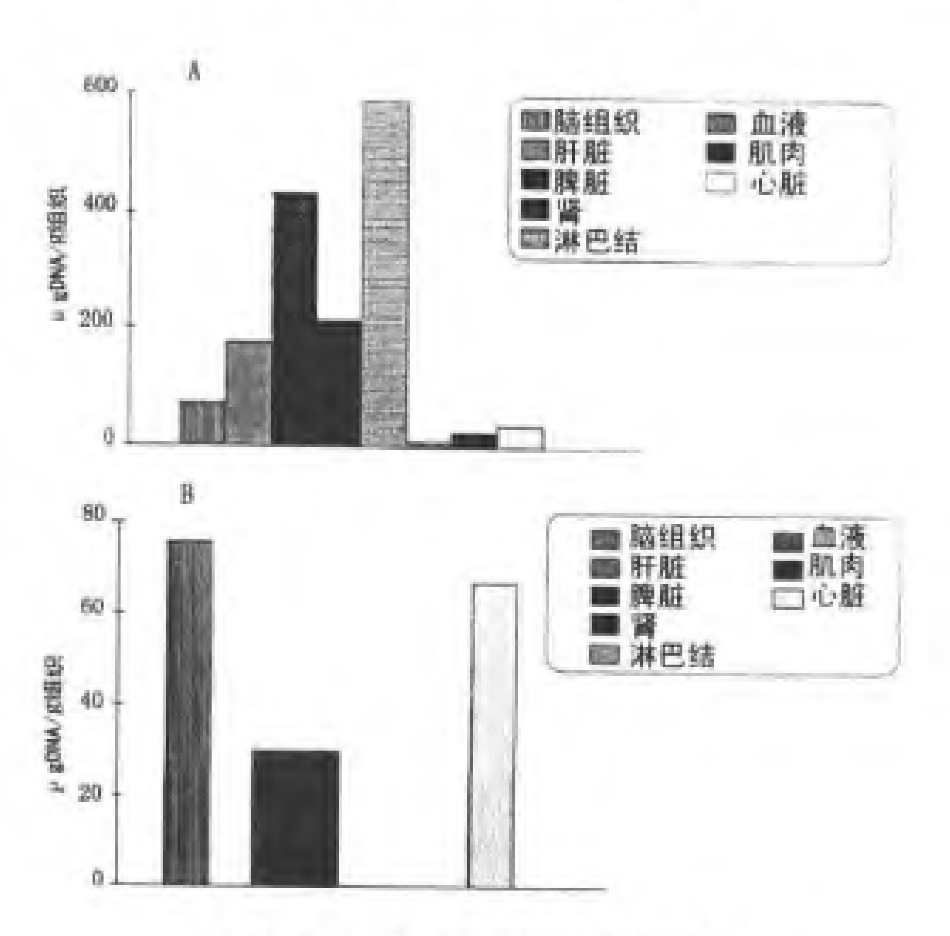


图 3-2　尸解后不同时间间隔的组织 DNA 产量

A：采集 7 天后；B：30 天后

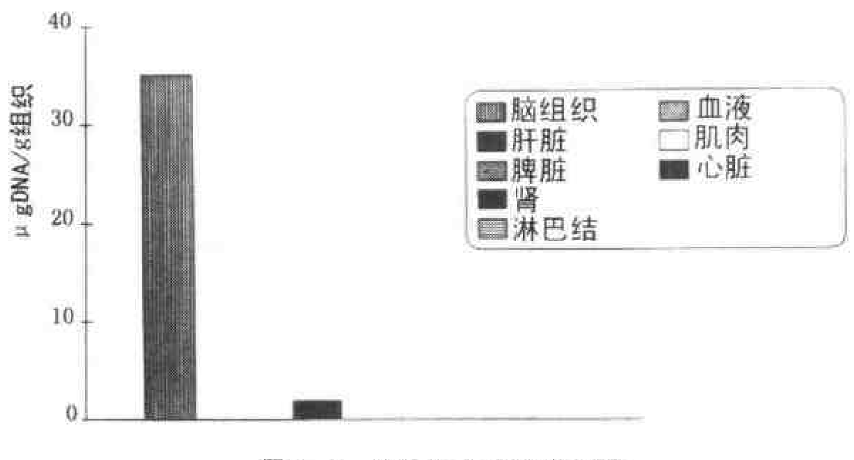


图 3-3 死后 85 天 DNA 提取量

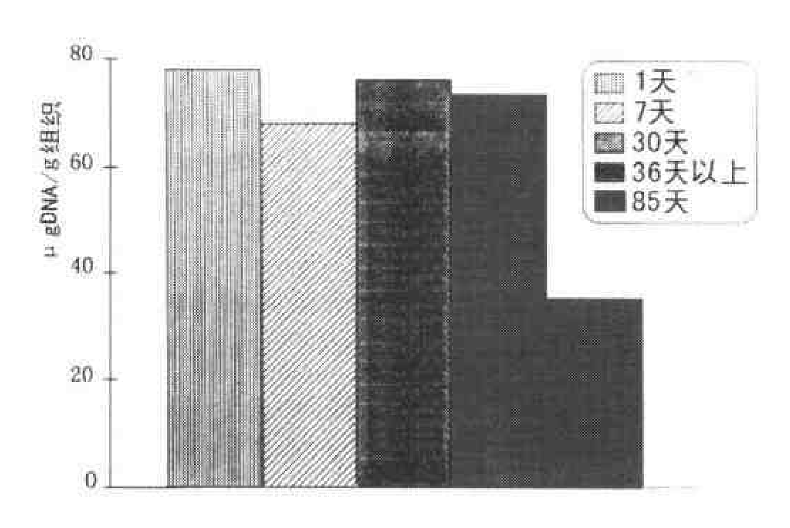


图 3-4 脑组织在不同保存时间的 DNA 产量

软骨由于钙化程度底，其中的 DNA 含量高且较易提取。如现场物证有软骨存在，应提取软骨送检，如肋软骨，会厌软骨等。

含有脱落表（上）皮细胞的其他材料如胃液、粪便斑、鼻涕、头皮屑也能提取到 DNA，由于摩擦、接触，指纹中也存在脱落细胞，也可以获得指纹的 DNA 分型结果。但是在一般情况下，很难从粪便斑等材料中获得足以进行标准基因分型的 DNA。

3.2 外界环境对生物检材 DNA 的影响

生物物证在采集、送人实验室检验前，一直处于无选择的外界环境中，会受到光、潮湿、高温及洗涤剂、细菌、灰尘、昆虫等的影响，可能对检验有以下影响：（1）干扰 DNA 提取过程，降低 DNA 产量，限制 DNA 分析；（2）存在的抑制剂在提取过程中被一起提取出来，抑制酶反应。（3）暴露于上述环境，可能损害 DNA，引起 DNA 的降解，影响 DNA 提取的质量或产量（表 3-2）。

表 3-2　引起 DNA 损伤的环境因素

环境因素	DNA 损伤可能形式	评注
微生物	链的剪切	高温（37℃）、高湿易长细菌，微生物分泌内切酶酶切 DNA
光	碱基修饰、链蛋白质交叉连接	短波照射
氧化剂	碱基修饰、链的剪切	
有机溶剂(乙醇、汽油)	一般没有影响	
酸	碱基修饰、链剪切、脱嘌呤	强酸或提高温度才能引起破坏
碱	链的解离、链的剪切	强碱（pH > 12）
温度	低温没有影响，高温引起 DNA 变性，变成单链	
时间	自发性的变性修饰、链的剪切和脱嘌呤	DNA 处理溶液状态，自发率很低，处于干燥状态自发率极低

3.2.1 DNA 降解

DNA 是存在于生物体内的大分子，在体内受到严格的控制，DNA 分子从离体那一刻起，就处于外界环境中，变化已经开始发生。一旦离开它本身的保护环境，DNA 分子很容易被损坏、发生断裂，变成小分子，称这种断裂为降解。DNA 断裂发生有二种情况：(1) 核酸酶（DNA 酶）的作用产生的断裂；(2) 各种外界因素引起的 DNA 断裂。前者有细菌和检材自身细胞内存在的核酸酶；后者有机械外力、紫外线照射、化学试剂如酸、碱、甲醛、漂白粉等强氧化剂因素。DNA 的降解影响 DNA 分型及结果的利用率，尤其对 DNA 指纹图和纹印图影响较严重。例如：正常 DNA 片段大小（也称为 DNA 分子量，用 MW 表示）在某一个 RFLP 位点是 20kb，当样品发生降解，DNA 分子平均大小可能只有 10kb 了，那么这样的 DNA 根本不可能再在原来 20kb 位置上检测到谱带，得到的 DNA 结果就不能反应样品原来的遗传信息。

3.2.2 影响 DNA 降解的因素

影响 DNA 降解的环境因素很多，主要有时间、温度、湿度、光（日光和紫外光），化学或生物物质。其中以阳光暴晒、紫外线照射、高温、潮湿保存、微生物和某些试剂等对 DNA 降解影响最大。Dwight E.Adams（1991）曾报道血斑在日光下暴露 8 周后，DNA 发生降解，在 RFLP 图谱上没有谱带；血斑在夏季白天平均最高温度 41℃和最低 23℃条件下暴露于日光 12 天，DNA 发生严重降解，得不到 RFLP 图谱，而在同样条件下存放于黑暗的血斑则可得到高分子 DNA；他还发现在 37℃下保存 4 年的血斑，只检测到小分子 DNA，没有高分子 DNA，RFLP 图谱检测不到谱带。表 3-3 为一日本学者研究的高温对样品 DNA 的 ABO 基因分型的影响结果，即使很短时间的高温环境也引起 DNA 降解，影响检验。

表 3-3 热处理对血斑样品 ABO 基因型分型的影响

加热条件	AO*	BO*	OO*	-
	O**	B**	O**	O**
	S1 S2 S3	S4 S5 S6	S4 S5 S6	S7 S8 S9
100℃ 10h	+ + +	+ + +	+ + +	+ + +
150℃ 1h	+ + -	+ + +	+ + +	+ + +
150℃ 2h	- - -	- - -	- - -	- - -
200℃ 1h	- - -	- - -	- - -	- - -

注*：血斑基因型；**：用特异引物扩增的目的等位基因

S1-S9 为样品编号；+：阳性结果；-：阴性结果，即没有扩增产物

潮湿环境不仅利于检材自身核酸酶活性的保持，而且潮湿环境利于细菌和霉菌生长，微生物分泌核酸酶，进一步加速 DNA 降解。一些洗涤剂或去污剂等化学试剂能够溶解细胞膜、核膜，使细胞通透性增加，DNA 从细胞中释放出来，并且使组蛋白与 DNA 分离，使 DNA 易受各种外力作用发生降解。

物证载体本身质地和所受污染情况不同，对于 DNA 的降解影响也不同。如地毯沾有很多灰尘或泥土和微生物，吸湿性又较好，潮湿的斑痕不易干燥，斑痕中 DNA 易发生降解。表 3-4 是从几种不同载体上提取的 DNA 分子情况。

表 3-4 载体材料对 DNA 分子的影响

载体类型	样品数	高分子 DNA	高分子 DNA 与部分降解 DNA 共存	降解 DNA	没有 DNA
碎屑	13	7	1	1	4
塑料袋	2	1	…	…	1
合成物	59	26	6	7	20
斜纹粗布	7	3	1	…	3
地毯	4	…	…	…	7
泥土	19	47	15	14	42

高温、潮湿环境最易引起 DNA 降解，而且最严重。事实上，在实际案件中，生物物证 DNA 的降解常常是上述多种因素的综合影响所致。

3.2.3 DNA 降解对分型的影响

1. 降低分析的成功率

DNA 降解降低 DNA 分析的成功率，但对各种方法分析成功率的影响是不一样的。对需要大片段的 RFLP 多态性分析影响较大，对只需小片段的 PCR，尤其是 STR 分析的影响较小。

2. DNA 图谱中的谱带减少

DNA 发生降解，分子变小，大片段可能会丢失，导致 DNA 图谱的谱带减少，如杂合

子只显示一条带。因此，当图谱表现为一条带时，要弄清样品是真正的纯合子（一条带），还是杂合子由于某种原因而导致二个等位基因（二条带）中的一个丢失，这就要求对标本的质量进行估计，以便准确解释结果。

3.2.4 采集和保存过程中 DNA 降解的阻止

现场发现的生物检材从提取、包装、保存，直至送到实验室的过程中，必须考虑环境因素对检材 DNA 降解的影响。原则上在干燥、低温条件下 DNA 降解较缓慢。因此现场提取生物检材后，应尽快晾干，低温条件下保存与送检。

3.3 物证的污染

含有 DNA 的生物组织和体液在离开生物体后，就可能被其他物质污染，从而影响检材 DNA 的分析结果。污染物不同，其影响 DNA 分析结果的原因和表现形式也不同。常见污染物大致可分为三类：（1）非生物物质；（2）非人的生物物质；（3）他人的 DNA。

3.3.1 非生物物质的污染

非生物物质类污染是最常见的污染。他们主要是染料、洗涤剂、灰尘和各种生活生产中包括一些化学制剂等在内的非生物物质。其中有些污染物在物证斑迹形成前就存在于形成斑迹的载体上，如布的染料、衣服上残留的洗涤剂，有的是物证斑迹形成后落在斑迹上的，还有的是在现场提取物证或实验室 DNA 提取过程中形成的。不同的污染物对 DNA 分析的影响是不一样的。有的影响比较重，例如深色布载体上的染料就能抑制 DNA 内切酶和 DNA 聚合酶的活性，DNA 分子不能被酶切，或者 PCR 扩增时得不到扩增产物，没有分型结果。

3.3.2 其他生物物质的污染

其他生物物质的污染包括各种动植物和微生物等。虽然法医学选择应用的 DNA 遗传标记一般是人特异的，能区别开人与动物，一般不影响最后结果的解释，但亦应注意防止这类物质的污染。一些动物 DNA 偶尔在人的 DNA 分析中也出现交叉反应。在进行 DNA 检验前，应进行种属检验。微生物的污染则影响比较大，需要特别予以注意。犯罪现场标本如血、精液是细菌和真菌生长的营养丰富环境，这些微生物分泌的物质如核酸酶等能降解样品中的人类 DNA。微生物的污染引起 DNA 严重降解，使分型难以获得结果。

3.3.3 人之间的 DNA 交叉污染

这种交叉污染可能发生在物证样品收集过程中或之后使物证以外的其他人的 DNA 落到物证上。主要有以下情况导致污染：（1）提取或接触物证的人没有带工作帽、口罩和手套，使头上、手上的人体成分脱落或粘附到物证的斑迹上，咳嗽或打喷嚏时唾沫、鼻涕落到物证上；（2）提取物证的器材污染有 DNA 或在提取各个物证之间没有清洗干净，造成交叉污染；（3）采集人员缺乏经验，把物理位置上不在一起的物证采放在一起，而这些物证可能含有来自不同个体的 DNA。（4）不同物证检材没有分开包装或包装和保存中样品隔离不够，造成物证间交叉污染或外流污染。一旦样品中污染有他人的 DNA，使物证的 DNA 检验受到干扰，在一个样品中检出二个或二个以上个体的 DNA，检验结果变得复杂，难以判别。

污染影响案件 DNA 分析的程度取决于所用检验方法的灵敏度和污染量。比如，应用

PCR 分型系统，由于样品中的 DNA 分子可以成千上百万倍地复制，因此，即使其污染 DNA 量很微，也能被检出。

3.4 物证采集

3.4.1 对物证采集的要求

影响 DNA 分型结果有三个因素：第一因素是 DNA 的量，虽然 PCR 方法很灵敏，但还是需要一定量的 DNA；第二个因素是 DNA 的质——DNA 样品的降解，前面已述 DNA 分子易受各种环境因子影响而降解；第三个因素是 DNA 样品的纯度，各种各样的污染会影响 DNA 分型。

不同的采集保存方法会在较大程度上影响以上三个因素，以致影响随后的 DNA 分析。因此物证采集的原则是：

1. 提取或接触物证的人要带工作帽、口罩和手套，杜绝用手直接接触物证；一定要防止现场勘查和提取物证人员的 DNA 污染；

2. 尽可能多地采集物证，以保证得到足够量的 DNA，满足分析；

3. 避免泥土、油脂等环境中其他物质的混入；

4. 所用采集器具要干净、无菌；要尽可能使用一次性器具；需要连续使用的在提取每件物证前都要彻底清洗、冲净；

5. 对于物理空间不在一起的物证应分开提取、分开包装；

6. 一旦物证提取后，要分别正确包装，做好标记。对于不会外流的物证，应当用纸袋包装；

7. 采取的物证样品，小件的要放入冰箱保存，大件的不能放入冰箱的要保存在干燥通风凉爽低温的环境下。斑痕类物证一定要先晾干；组织类物证最好冰冻保存；

8. 尽快送检。

3.4.2 物证提取的基本方法

附有 DNA 的生物物证载体各种各样，为适合随后的实验室分析，可酌情采取以下二种基本方法中的一种进行：(1) 连同载体一并采集。(2) 将斑痕从载体上转移到更合适或更易处理的物体上。

第一种方法比较好，样品不会受到损失，受污染的危险性小，操作简单。提取人员用正确方法将整件物证提取、包装好，径直送到实验室分析。这种方法适合于易于包装和运送的像衣物或其他比较小的物体。也适合那些可以切割的大件物体，如门框或窗帘，可以将有斑痕及斑痕附近的客体一起完全切割下来送检。将斑痕从客体上转移下来的工作留给检验人员去处理较妥，因为他们知道如何做更有利于检验分析。

对于独立的、没有附着在载体上的物证，如骨骼、毛发、指甲等，亦按此种方法提取。

第二种方法是将生物物质转移到更好的载体上，比如，将大块石头、水泥地面、沥青地上的斑痕转移到纱布上。这可以用干净的手术刀刮取样品或用胶带粘取，或用蒸馏水、或生理盐水、或缓冲液蘸湿的纱布将斑痕擦拭下来。用刀刮取或胶带粘取不需要弄湿样品，不会带进潮气导致 DNA 降解，但是难于完全将物质转移到载体上，可能损失部分样品或在擦刮过程中不小心损失了样品。采用浸湿擦拭法提取的斑痕物证，要立即晾干保存。

如果斑痕在刀具、铁棍或其他铁制品表面，即使暂时不进行 DNA 检验，也要尽早将它们转移下来。因为铁器容易生锈，生锈后导致无法辨认斑迹的残留部位。

物证采集的标准程序要求在斑痕附近采集一个或多个没有斑痕的样品，作为检验时的对照。

3.4.3 注意物证存在的多样性，广泛提取各种 DNA 物证

有试验证明人体只要接触物体 5 秒钟，就有细胞遗留在物体上。由于紧密、频繁接触，可在各种物体上留存人体细胞，例如：在帽子、面罩、围巾上可能遗留有使用者的头面部脱落的细胞、毛发；在烟头、口香糖中、各种物体的咬痕上可能留有口腔上皮细胞；在缢勒案件的绳索上可能留有皮肤和血细胞；在受害者抓罪犯时，受害人的指甲缝内可能留下罪犯的细胞；在枪伤案件中的弹头上可能留有伤者的细胞；在贴身衣服、手表、眼镜、杯子、电话耳机、用过的铅笔及牙刷等各种物体可能留下使用者的表皮或上皮细胞。

在特殊案件中，应注意搜集在案件过程中可能留有的这类 DNA 物证。这些物证通过 DNA 分析，对于查明案件中的某些情况，常常可以提供科学的证据。

3.5 物证的标记与登记

物证在提取采集过程中应逐件进行标记和登记（采集日期、地点、名称、内含物、状况和数量）。像 DNA 分析这种结果具有较高准确性、科学性与客观性，鉴定结果对于判案具有重要作用的物证决不能有差错。一旦样品在分析前发生了任何差错或混入了其他样品，那么所得的结果就会变得不可靠或难以解释。为杜绝此类事情发生，使物证鉴定符合法律要求，物证的采集应严格遵照有关规定进行。在物证的采集前后都要注意笔录，要标记好并登记清楚。在现场，在搬动和提取物证前首先要照相，记录物证的位置与状况，注明物证与现场其他物体的空间距离，采集完后要在包装上标记清楚。

3.6 物证的保存

进行 DNA 检验的生物物证一旦收集后，必须保持在干燥状态，潮湿的物证应室温晾干，低温保存，以减少 DNA 的降解，不要暴露于高温、潮湿、或阳光直射的环境。人体组织不能加福尔马林等固定，应采取冷冻保存。

3.7 各类物证的采集与保存

所有需要 DNA 分析的生物物证都应遵循上述物证的采集和保存原则，此外针对一些物证的特点，对这些物证采集和保存再作进步阐述。

3.7.1 血液和血斑

1. 个体血液

（1）由医务人员采个体血。

（2）使用无菌一次性的注射器收集 1～5ml 的血液，放入含有 EDTA 抗凝剂的试管中，充分混匀，加封。EDTA 是 DNA 的最好抗凝保存剂。或者将抽取的血液直接涂抹在无菌纱布或干净滤纸上，晾干，制成血痕备用，也可采集指血，制成 $1cm^2$ 以上的血痕。如果要进行 DNA 指纹分析，以液体血液送检为最佳。

(3) 每个试管应标明日期、时间、嫌疑人的姓名、地点、采血员的姓名、采集的数量、案件名称编号和物证编号。

(4) 血液样品应置于4℃冰箱保存，制成血斑的，则在室温晾干，并应尽快送检。

2. 现场的液体血

(1) 液体血应该使用干净的（最好是灭菌的）注射器或一次性吸管，将其吸至一个清洁的（最好是灭菌的）试管内。

(2) 凝血块可以使用清洁的药匙将其移至清洁的试管内。

(3) 可以使用清洁的纱布吸取液体血或凝血块。不要只吸取血清部分。血痕在包装和送往实验室前应晾干（不要阳光直射）。

(4) 标明样品案件名称编号、物证编号、日期、时间、采集物证的位置，以及物证采集员的姓名。

(5) 液体血液样品应使用 EDTA 抗凝剂，冰箱保存。

3. 雪上、冰上和水中的液体血

(1) 雪上和水中发现的血液应立即提取，以免被进一步稀释。

(2) 将那些最浓的样品尽可能地搜集到清洁、适合的容器内。采集时应尽可能避免杂质污染。

4. 衣服上的未干血斑

(1) 将带有未干血斑的衣服放在一个清洁的表面，晾干（避免阳光直射），不能用电吹风等热力加速干燥。

(2) 未干的血斑及载体本身潮湿的不能密封在不透气的密封袋或塑料袋内；否则样品保持在潮湿状态下，易于细菌繁殖，会使样品 DNA 降解。

(3) 用干净的纸将它们包装起来，或装入纸袋内。

(4) 不能抖动带血斑的衣服，以防引起血痂的脱落转移。

5. 物品上的未干血斑

(1) 带有未干血斑的小件物品应晾干，晾干后整件收集并包装。

(2) 对于不宜从现场移出的大件物品存在的未干血斑，将其转移到干净的棉纱布上，晾干后装入纸带内。

6. 大的或不可移动物品上的干血斑，如墙壁或混凝土上的干血斑

(1) 对血斑的形状必须做笔录、照相并且画图。

(2) 将血斑直接刮到一张清洁的纸上；或用高粘质胶带将每一个斑点从其所在的物体表面上粘下来；或用蒸馏水湿润的清洁的纱布或拭子将样品擦洗下来；或用一个小吸管吸少量蒸馏水于斑痕上，反复吸排多次，将斑点从物体上洗下来，然后将样品吸到一个清洁的试管内。

(3) 转移到纱布或拭子的血要晾干，再将其放在专门的纸袋。

(4) 每一件物证都要收集没有斑痕的部分，作为“基质对照”送验。

7. 可以切割的物体，如地毯或室内装潢材料上的干血斑

(1) 详细笔录、照相斑点情况。

(2) 用一个清洁的刃器将带有血斑部分的物证材料切割下来。

(3) 对每个切割下来的物证都应该分别包装和标记。

8. 小的飞溅的血斑

小的飞溅的血点常常很难从物体的表面上转移下来。在做了适当的记录以后，可以用胶带或湿润的消毒棉纱线将样品提取下来，应避免从周围区域提取脏的或其他污染物。

9. 交通肇事案件中汽车上的血斑

(1) 全面检查汽车的外表面，确定有无毛发、组织、血液、以及其他的痕迹物证。所有的检查发现都要一一记录。

(2) 用消毒的利器（如解剖刀、凿子、螺丝刀）将斑点刮入一个清洁的纸袋内。

(3) 如果刮削的部分不令人满意，也可以使用湿纱布转移提取血斑，然后晾干包装。

(4) 如需要进行指纹显现，宜用对物证 DNA 检验无害的显现方法，如铝粉法显现指纹。否则应在显现指纹前提取血斑。

10. 尸体上的血斑

(1) 尸体上的血斑在提取前应仔细记录。

(2) 应注意血斑的位置、大小、数量、形状和类型。

(3) 由于尸体上血斑易于脱落，因此在移走尸体以前将血斑提取下来。

(4) 在收集尸体上的血斑时，为了减少附带尸体本身的皮肤细胞，应尽可能轻轻地将斑点取下。如使用胶带粘取或拭子轻轻地擦拭斑痕。

(5) 指甲下的血斑应该使用清洁的牙签刮下。用清洁的指甲剪将指甲剪下并收集起来。剪切指甲时应小心，不要带死者本人的血液和组织。每个指甲和牙签均应分开包装。

3.7.2 精液和精斑

1. 现场上发现的液体精液物证

(1) 精液物证应该照相、录像和画图记录。

(2) 用一个清洁的注射器或一次性吸管将液体精液吸至一个灭菌的试管内。

(3) 也可将液体精液吸取到干净的纱布或棉球上。然后晾干、包装、加封和标记。如果有条件，将样品保存在冰箱内。

2. 可移动物体上精斑

(1) 裤子、衣服、床单、枕头、纸张，以及其他可移动物体上的精斑整件进行收集。

(2) 如果物品上的斑点是湿的，那么在收集包装以前必须将其彻底晾干。

(3) 每一件物证都必须单独用一张清洁的纸包装起来，并做好标记。

3. 可以切割的大的物体上的精斑

(1) 大的物体，如地毯、床垫、室内装潢材料上的物证样品可以将其切割下来。

(2) 如前所述，所有这些物证都要做好采证记录。

(3) 用清洁的解剖刀将精斑切在一张清洁的纸上，然后折叠包起来。

(4) 解剖刀在使用中间应该用漂白粉和酒精彻底洗净，以避免任何可能的交叉污染。

(5) 每一个物证折叠小包都应放入一个单独的小包装袋内。

4. 不可移动的、非吸湿性物体表面上的精斑

(1) 不可移动的具有非吸湿性表面的物体，如地板、柜台、金属表面的样品等。

(2) 如前述，提取精斑物证必须做好采证记录。

(3) 用干净的解剖刀将精斑刮到干净纸上，然后将其折叠包起来。也可以用胶带粘下来或用棉球拭子将斑痕擦拭下来。

(4) 解剖刀在使用中间必须彻底清洗。

(5) 每个折叠包放入一个单独的包装袋内，所有的包装袋都要封好。

5. 从性犯罪案件受害者身上提取精液物证

(1) 性犯罪案件受害者的物证的采集应在医院、门诊室进行。

(2) 应按前述已制定的程序进行物证（如衣服等）的收集。

(3) 根据被强暴案情，用无菌纱布或棉球反复擦拭阴道、口腔（口交案例）和肛门（鸡奸）。纱布或棉球的面积千万不要过大，否则因擦拭物面积过大，精液斑不集中，影响下一步检验。提取的擦拭物要晾干保存。

3.7.3 组织、器官和骨骼

1. 新鲜组织、器官和骨骼

(1) 每件物证都要采用各种方法进行描述，包括笔录、照相或录像做采证记录。

(2) 每个物证都要使用一个清洁的器械进行收集。所使用的镊子在使用之中，如前述，要彻底地进行清洗。对腐败尸体，要取深层肌肉或肋软骨。放在一个清洁的可固定的容器内。

(3) 密封后，做好标记并放冰箱或冷处保存。

2. 陈旧组织、器官和骨骼

(1) 在收集前应该对每个物证进行拍照和画图。详细记录大小、形状、类型及其和其它物证相互之间的空间关系。

(2) 每个物证都应使用清洁的镊子提取。对仍然连在一起的物证应该一块儿提取。收集时不应将其分开。

(3) 骨骼尽量提取长骨，而且不要人为地将长骨切成两截。否则，骨内 DNA 容易被污染。

3.7.4 唾液、尿液和其他体液

1. 液体样品

(1) 液体的唾液或尿应尽快装入一个清洁的消毒的容器内（塑料的或玻璃的标本杯）。

(2) 每个容器上都要加封并做好标记。

2. 斑迹

(1) 唾液斑、尿斑和其他体液斑可以整件物证收集，也可以将其从载体上刮下或剪切下，予以收集。

(2) 物证必须彻底干燥并且放在一个清洁的纸质的容器内。刮削物和剪切物应收集在一张清洁的纸上并折叠包好。然后再将其放入第二个纸包装袋内，加封并做好标记。

3.7.5 毛发

(1) 用清洁的镊子夹取毛发。

(2) 不在一起的毛发应分别搜集、包装，做好标记。

(3) 在收集过程中必须小心操作，不要损坏可能存有的毛根组织。

(4) 与血液、组织、或其他体液混合存在的毛发，提取时应小心处理。每个物证都要

单独放在一个清洁的容器内，加封，并做好标记。

（5）如果毛发与体液混在一起，应晾干后包装送检。

3.7.6 强奸致孕案件的物证

受害人被强奸致孕，如果已人工引产，可用干净无菌的手术刀切开胎儿颅骨，用干净无菌的药勺或滴管取引产胎儿脑组织，放入干净试管或烧杯内，或取胎儿肌肉，登记受害人姓名、采集时间、地点和采集人等，密封，冷冻保存，送检。如未人工流产（1～3个月胎儿）可在医院采集绒毛。如要人工流产，一定在流产前告知医院妇科彻底清洗吸引泵，不能留有其他人的组织，而且吸引时不能将胚胎弄得太碎，以免分不清是胚胎还是母体组织。

3.8 物证的送检

提取的物证要尽快送往实验室进行检验。在送检现场物证的同时，采集与案件有关人员的样品并一同送检。有关人员的样品一般均采集其血液，或制成血痕，血液采集方法同前。在物证送往实验室的过程中也要注意避免物证 DNA 发生降解，为此在送检过程中要尽量避免物证经受高温、高湿、日光照射以及过长的运输时间，鲜血要求放在冰壶中送检。

不同案件需要收集的有关人员样本如下：

1. 现场物证为精斑的案件：要求采集受害人、有关犯罪嫌疑人的血液（痕）；如果精斑的载体上有可能留有其他人的精斑如受害人丈夫等，必须采集他们的血液（痕），一同送检。

2. 强奸致孕的亲子鉴定案件：采集受害人血液、受害人所怀胎儿组织或绒毛（未人工流产）与有关嫌疑人血液。

3. 确定身源的亲子鉴定案件：除物证外，还需采集嫌疑父母血液、或妻儿（或丈夫和孩子），如因某种原因不能采集到其嫌疑父母血液、或妻儿（或丈夫和孩子）样品，可采集嫌疑人的同胞血液，并尽可能多地采集所有同胞血液，或同一父系的或同一母系的亲属血液送检。

4. 个体识别类的案件：现场物证以及比对样品，如嫌疑人、受害人血等。

3.9 物证 DNA 分析结果的预测

一个物证进行 DNA 分型以前，有时需要做预实验来评估生物物质的种类，假如分析的是蕃茄酱或鞋油，DNA 全套分析只是证明没有结果，未免有点太浪费。因此在现场采集物证前或实验室提取 DNA 前就要进行血、精液或唾液等各种液体的预试验。

通过预实验，一旦确定样品为生物物质，就应进行某些基本实验来确定样品中 DNA 状况，了解 DNA 的质量（降解情况），有多少总 DNA 存在，在总 DNA 中有多少是人类 DNA。琼脂糖凝胶电泳可以了解有多少总 DNA 存在以及降解情况，通过斑点杂交可以确定总 DNA 中人 DNA 的含量。弄清 DNA 情况对决定采用哪种分析手段与对结果解释是非常关键的。

3.9.1 RFLP

RFLP 分析需要一定量的高分子量（HMW）DNA，HMW DNA 指平均片段大小在 20～

23kb 的 DNA。当用 RFLP 分析一个严重降解的 DNA 分子时，可能会有大分子区域的谱带丢失的危险性，比如，一个本来二条带的图谱可能变成了一条带的图谱，或样品的第二个供体的较大分子的谱带可能丢失了。预先不用琼脂糖凝胶电泳来检测样品的降解程度，可能会导致假排除。

另外，RELP 能成功分析的 DNA 最小量与实验室有关，也与探针标记方法有关。一般说来，多基因座 RFLP 的 DNA 量为 1μg～8μg（非同位素标记）或 1μg～5μg（同位素^{32}p 标记），单基因座 RFLP 则为 10～50ng，换句话说，必须有 10～50ng 以上的高分子 DNA 才能得到理想的 RFLP 图谱。

3.9.2 PCR

许多 PCR 系统不像 RFLP 那样需要完整的 DNA（未降解），而且需要的 DNA 量也更少，有些系统如 STR，只要几百 bp 大小的片段即能得到理想的分析结果。PCR 系统的主要优点是它的高灵敏度。有些系统只需 0.2～0.5ng（比 RFLP 少 100 倍）样品 DNA，即相当于几百个精子或针头大小点的一滴血中的 DNA 就能得到基因分型。

除了 DNA 量和质外，污染的存在与否也是评价物证的内容之一。斑迹载体材料（如树片，泥土，蓝色牛仔裤、玻璃）、斑迹形成前或过程中或形成后加入的化学物（如油脂、精子灭活剂，洗涤剂，苏打等）都会影响样品的提取和分析。了解样品的情况将有助于决定下步采用什么方法分析最合适。

例 1：已知蓝色斜纹布含有与 DNA 结合的染料，抑制 RFLP 酶切和 PCR 扩增，目前可以用以 silica 为基础的纯化试剂盒除去染料，可以进行全套分析，因此当分析人员手中的斑痕在蓝色牛仔裤或深色棉布上时，就要考虑采用纯化程序。

例 2：从土中采集的物证，由于土壤中含有许多微生物，容易被降解，含有大量非人类 DNA，面对这种物证，分析人员就要弄清样品中含有多少人 DNA、多少非人 DNA 和人 DNA 降解程度。在许多情况下，不论其斑迹有多大，分析人员只能用 PCR 方法进行分析。

第四章 DNA 的提取

4.1 DNA 提取方法

DNA 存在于细胞内，在进行 DNA 检验前，需进行提取，即将 DNA 从细胞中提取出来并与细胞其他成分分开。现场提取的血斑、精斑和唾液斑等体液斑常留存于各种载体上。因此，生物检材的 DNA 提取也包括与样品中存在的其他非生物物质分离开来的过程（图 4-1）。任何残留的非 DNA 物质都会影响随后的分析，其影响方式有两种：第一，影响酶的作用，在检验过程中应用的 DNA 酶试剂需要在特定的环境下才能对 DNA 准确地有效地发挥作用。自然规定这些酶作用的对象只是 DNA，当没有外源其他物质存在时它们处于最佳工作状态；第二，一些外部物质可能会引起 DNA 降解，这种降解甚至在分析过程中仍将继续存在，直至这些物质被去除后才能停止。因此在进行 DNA 提取时总是要选择那些能尽可能彻底地、快速高效地去除任何可能影响 DNA 分析的有害物质的方法。

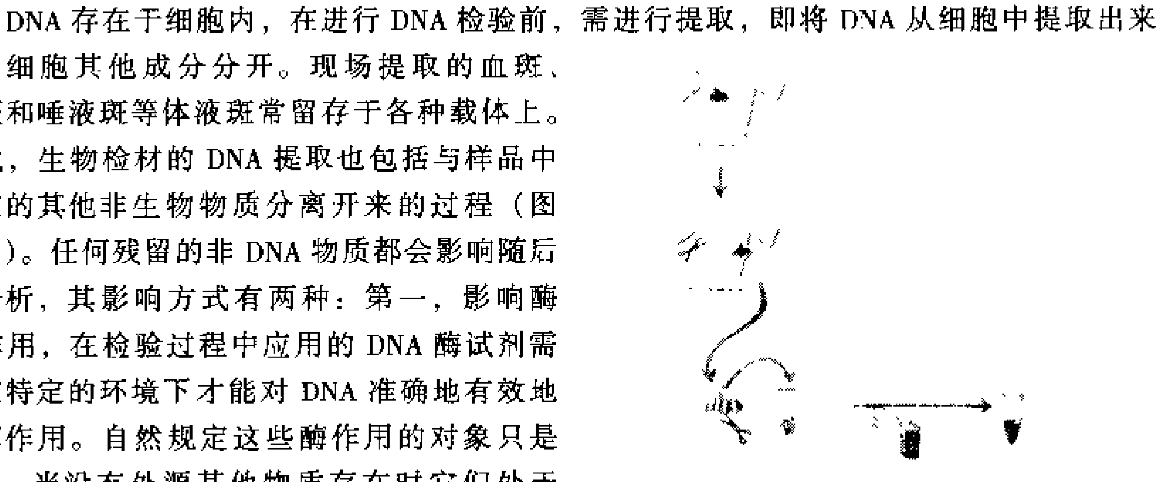


图 4-1 斑痕检材 DNA 提取示意图

核 DNA 提取的原理是：首先用低渗或高渗法裂解细胞，通过离心收集细胞核，核内 DNA 通过加入 SDS（十二烷基磺酸钠，一种去污剂）或加热煮沸法将核膜溶解、破裂，细胞核内的 DNA 从核内释放；加入蛋白酶 K，水解与 DNA 结合的蛋白质，使 DNA 游离在溶液中。随后用不同方法进行抽提，除去杂质，收集 DNA。

DNA 提取程序，根据生物物证种类（如血、精液、唾液和毛发等）、物证量和细胞种类的不同而有所些不同。可以通过肉眼观察，显微镜检查和预实验等决定选用哪一种提取方法。目前常用的方法有：（1）有机提取法；（2）Chelex-100 提取法；（3）无机提取法；（4）其他提取方法。

4.1.1 有机提取法

在多数情况下，采用有机提取法，即经典的饱和酚-氯仿法。它利用有机的蛋白质变性剂酚、氯仿交替抽提，使蛋白质变性，有效地去除蛋白质等杂质。有机提取法能获得较多的大片段 DNA，比 Chelex-100 提取方法能更彻底地纯化 DNA。它利用 SDS 将细胞膜裂解，在蛋白酶 K、EDTA 和去污剂的存在下消化蛋白质成多肽或小肽分子，然后用等体积饱和酚抽提除去蛋白质，酚-氯仿抽提去除未消化的蛋白质，再用氯仿抽提除去 DNA 溶液中的酚，最后用 2.5 倍体积的无水乙醇在 0.3mol/L NaCl 存在下沉淀 DNA，浓缩回收 DNA，70%乙醇洗去 DNA 沉淀中的盐，真空干燥后用 TE 溶解 DNA。关于 DNA 的沉淀，将

在 DNA 浓缩部分介绍。

1. 提取操作

(1) 加入 1.5ml EDTA - Na_2 提取液，悬浮各种来源处理后的细胞（核）沉淀，并反复吹打，使沉淀分散开来。

(2) 加入 1/10 体积的 10%SDS，使最终浓度为 1%，混匀。(SDS 作用：溶解核膜，使 DNA 从核中释放出来)

(3) 加入 1/20 体积 2mg/ml 蛋白酶 K (protease K，PK)，使最终浓度为 100μg/ml，混匀，封口，于 37℃水浴保温 4h (PK 作用：水解蛋白质)。如在 56℃下保温，时间可减少 1 ~ 2h。EDTA - Na_2 提取液也可用 NaAC 代替。

(4) 向溶液中加入等体积的饱和酚，以上下颠倒试管方式混匀液体，将管内溶液混合成乳状液。3 000 ~ 5 000r/min 下离心 5min，此时溶液分成三层，DNA 在无色的上层水层（水相）中，中间白色混浊层（有机相）为变性蛋白质。如果 PK 水解蛋白质彻底，此层物质就会变得少些。最下层带褐或黄色的是酚层。

(5) 转移上层水溶液到一干净试管中。如为了增加 DNA 回收量，可以将有机相取出，加入等体积 TE 溶液 (pH7.8)，充分混匀后，离心，再吸取上层水相，与第一次抽提的水相合并。

(6) 加入 1/2 体积的酚和 1/2 体积的氯仿，混匀。3 000 ~ 5 000r/min 下离心 5min。

(7) 如果水相层和有机层的界面不甚清楚，说明其中含蛋白质较高，可采用多次酚/氯仿溶液抽提，并可适当加大离心速度或延长离心时间。此步中可以加少许异戊醇，目的是减少在蛋白质变性操作过程中产生的气泡。

(8) 转移上层水溶液到另一干净管中，加入等体积的氯仿混匀，3 000 ~ 5 000r/min 离心 5min。转移上层水溶液到另一干净管中，加入 1/10 体积的 3mol/L NaCl，轻轻混匀，加入 2.5 倍溶液体积的冰冷（-20℃保存的）无水乙醇，混匀，沉淀 DNA。

(9) 收集沉淀，用 70%乙醇洗去 DNA 中的盐，5 000r/min 离心 5min，除去液体，真空抽干 DNA 5min，至 DNA 成无色透明状（注：干燥时间随使用真空度变化，或长或短）。

(10) 加适量 (50 ~ 100μl) TE 缓冲液溶解 DNA，溶解后的 DNA 溶液于 4℃保存。

用 Microcon100 超滤管旋转离心代替乙醇沉淀可获得大量 DNA，如果溶液中 DNA 含量较低，乙醇沉淀法回收率低。

有机提取法用来获得 50kb 以上的大分子 DNA，在加 SDS 后所有过程中，要避免剧烈振荡，在移取 DNA 过程中，可以用弯头吸管或将加样器的吸头的尖部剪去来移取 DNA 溶液，因为 DNA 溶液较粘，加样器吸头的尖部孔径较小，用其吸取时液面的局部压力较大，易破坏溶液分层界面。操作中要带手套，防止皮肤、眼睛接触酚。

4.1.2 Chelex - 100 **提取法**

当检材量较少时，或者基因分型只需要少量的 DNA 时，可以用 Chelex - 100 提取方法提取 DNA。Chelex - 100 是一种化学螯合树脂，由苯乙烯、二乙烯苯共聚体组成。含有成对的亚氨基二乙酸盐离子，螯合多价金属离子，尤其是选择性螯合二价离子，比普通离子交换剂具有更高的金属离子选择性和较强的结合力，能结合许多可能影响下一步分析的其他外源物质。通过离心除去 Chelex 颗粒，使这些与 Chelex - 100 结合的物质与 DNA 分离，

防止结合到 Chelex 中的抑制剂或杂质带到 PCR 反应中，影响下一步的 DNA 分析，并通过结合金属离子，防止 DNA 降解。EDTA 虽也具有螯合金属离子作用，但它结合的金属离子仍留在溶液中，无法与 DNA 分开，影响 DNA 分析。这就是为什么 Chelex - 100 不同于 EDTA 的原因。

Chelex - 100 提取法基本步骤：用无菌蒸馏水裂解细胞或其他方式得到细胞核，用 5% Chelex - 100 的悬浮液悬浮细胞核，在 56℃保温 30min 以上，然后 100℃保温 8min，煮沸可以裂解细胞核，释放 DNA，同时使蛋白质变性成为不溶物，最后离心除去与金属离子结合的 Chelex 颗粒和其他非 DNA 成分（如蛋白质）。DNA 留在上清液中，可直接用于随后的分析。56℃保温时向溶液中加蛋白酶 K，可增加裂解的效率。对于血液（斑），56℃保温时间的长短对提取影响不大，保温 30min 与过夜没有明显差异。

取 Chelex - 100 悬浮液时，应该将颗粒平均分散开来，在加样前将悬浮液充分摇动，并且用 1ml 吸头或前面较粗的吸头取 Chelex - 100。

如果 Chelex - 100 提取的 DNA 在放置一段时间后还要分析，应当把上清液与 Chelex - 100 颗粒分开，保存在 - 20℃。Chelex - 100 会随时间的延长被破坏，释放出以前结合的抑制剂或树脂本身结合 Mg^{2+} 抑制 PCR 反应。

Chelex - 100 结合效果依赖于悬浮液 pH 值。5% Chelex - 100 的 pH 值在 10.0 ~ 11.0 为好。如果发现悬浮液的 pH 值不在此范围内，则应废弃。5% Chelex - 100 悬浮液最好现用现配，因为 Chelex - 100 在数小时后螯合力有下降的趋向。

在 Chelex 方法中，细胞 DNA 一直在同一管中进行操作，不涉及转移，因此检材的损失率低，是一个十分简单、快速的方法。但底物、核膜及其他试剂一起保留下来，DNA 纯度比有机溶剂提取法低。Chelex - 100 方法将 DNA 双螺旋打开，不能用像 EB 染料插入方法进行定量分析，只能用杂交方法进行定量。Chelex - 100 法只适用于 PCR 分析，而不适合于 RFLP 分析，而且对于目的片段大于 500bp 的基因座的扩增成功率低。因为用于 RFLP 分析的 DNA 必须是双链大分子 DNA，100℃高温煮沸 8min 和快速振荡使 DNA 片段受到损害，片段断裂。

4.1.3 无机提取法

无机提取法是利用 6mol/L NaCl 或醋酸钠沉淀蛋白质代替有机酚 - 氯仿抽提步骤去除蛋白质，其他过程同有机溶剂提取法。大致步骤是：在消化好的 DNA 溶液中加入 3 倍体积 6mol/L NaCl，剧烈振荡混合，5 000r/min 下离心 15min，去沉淀；上清液用无水乙醇沉淀 DNA。此法操作简单，步骤少，不使用有毒害的化学试剂酚和氯仿，但是得到的 DNA 溶液盐浓度较高。

4.1.4 差异裂解提取法

在法医物证检验中现场样品经常涉及一些特殊情况，比如样品中含有精子和其他类型的细胞，常见是阴道来源的细胞，也就是说混合精斑。在法医物证检验中发现的包括唾液、皮肤、口腔、阴道细胞以及在尿液和排泄物中的自体细胞大都为上皮细胞（epithelial cell，简称 e cell，E 细胞）。这些上皮细胞和精子细胞的性质不同，精子细胞膜和核膜含有大量硫醇蛋白，含有丰富的二硫键，比其他细胞如血、唾液或阴道上皮细胞更难打开，需要用特殊的化学试剂如二硫苏糖醇（DTT）打开二硫键。因此可以利用精子这种特性，在

DNA 纯化前使精子 DNA 与其他细胞 DNA 分开，这样，可将罪犯和受害人的 DNA 分析分开进行，从而简化了最后结果的解释。这种用于分离混合样品中精子细胞和表皮细胞的程序叫差异裂解提取。

差异裂解提取和有机溶剂提取程序不同，它通过把检材样品浸泡在温和的溶液中，首先将细胞（包括表皮细胞和精子）与底物（载体）分开，让样品在只能裂解上皮细胞的化学试剂中温浴，然后离心，将含有裂解上皮细胞部分的溶液移人另一管中。这部分来自于裂解上皮细胞部分的溶液，再经有机溶剂提取法进行 DNA 提取，可用于上皮细胞部分 DNA 检验。而对检材样品中的精子细胞则另加一些试剂（DTT）帮助精子细胞裂解，这部分来自于裂解的精子细胞溶液，叫精子部分。用有机溶剂提取法等提取 DNA，进行精子细胞的 DNA 分析。

在实际案件物证中，由于检材样品的条件不同，上皮细胞和精子细胞的分离可能并不都是非常完全的。例如，样品条件差，一些精子细胞在上皮细胞裂解程序中已经发生了裂解，提前释放 DNA，使精子的 DNA 出现在上皮细胞部分的溶液中。相反的，在有许多上皮细胞而只有几个精子的混合物中，一些上皮细胞在第一步裂解中未被裂解，上皮细胞可能留在精子中，在最后的精子部分检验中检出上皮细胞 DNA。在这两种情况下得到的结果比较复杂，表现为混合样品，结果有时难予解释。

4.1.5 其他方法

1. 磁性珠提取法

该方法利用磁性硅胶吸附白细胞，用细胞裂解液裂解细胞，从细胞中游离出来的 DNA 分子被吸附到磁性颗粒表面，蛋白质等分子不被吸附而留在溶液中。在磁场作用下，磁性颗粒与液体分开，回收颗粒，再用纯水或 TE，洗脱吸附的 DNA（图 4－2）。磁珠法不需要离心，操作简单，而且在同一个反应管内提取，没有样品的损失，更适合于自动化提取。与 Chelex－100 提取法比较，磁珠法回收率高，即使是 0.5μl 血仍能得到 DNA 分型，而用 Chelex－100 提取同样量的血 DNA 有时不能得到分型。

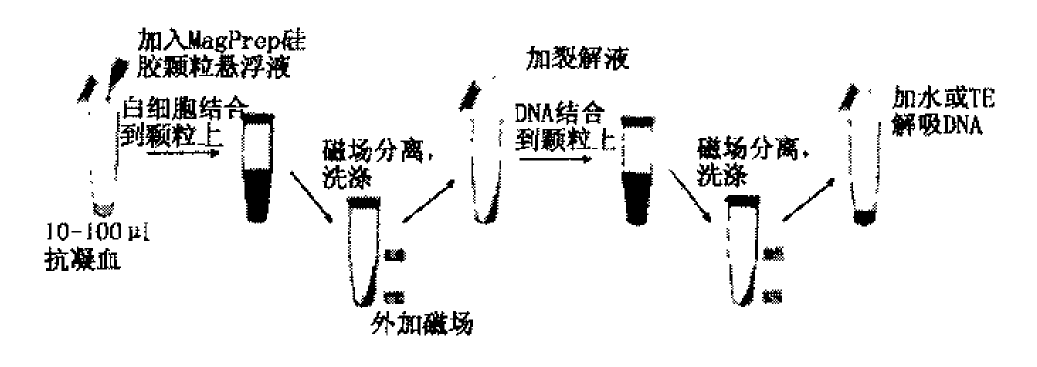


图 4－2 磁性珠法提取 DNA

方法：取 10μl 抗凝血，加入 200μl 磁珠悬浮溶液，室温放置 5min，然后将溶液放到磁场中 1min，吸去溶液，加人 200μl 裂解液，室温裂解 5min，然后将溶液放到磁场中 1min，吸去溶液，加人 200μl 洗涤和洗脱液洗，重复洗一次，此步要求不要搅散磁珠；最后加 50μl 洗涤和洗脱液将 DNA 从磁珠中回收出来，此步要求充分悬浮溶液，使 DNA 得到充分

释放。如果提取血斑 DNA，在最后加 50μl 洗涤和洗脱液后，将之在 95℃保温 8min，然后回收 DNA。

2.NaI 提取法

样品放入 1.5ml 的离心管中，加 300μl 消化液（6mol/L NaI－13mmol/L EDTA－0.5%月桂酰钠－10μg 糖原－26mmol/L Tris－HCl－29nmol/L DTT，pH8.0），混匀，55℃保温 4h。加等体积 80%异丙醇，温和混匀 15min，12 000r/min 离心 10min 沉淀 DNA。加 1ml 40%异丙醇，充分振荡，将 DNA 沉淀物弹起来，12 000r/min 离心 10min 回收 DNA。真空干燥，加 20μl 蒸馏水溶解 DNA。

3.Dynabeads DNA direct kit

加 3ml 白细胞裂解液（10mmol/L Tris－HCl，pH7.6，10mmol/L EDTA Na_2，50mmol/L NaCl），42℃保温 3～4 小时后，加 15μl 0.5% NP－40，50μl PK（2mg/ml），42℃保温过夜。加 200μl 磁性珠，室温放置 5～10min 后，放到磁场 90～120s，去上清液，在留下的珠－DNA 复合物内加 200μl 试剂盒带的洗涤缓冲液，放到磁场 90～120s，去上清液，最后将 DNA 溶于 60μl　TE 缓冲液中，65℃保温 5min 以分离复合物，直接收集上清液，－20℃保存或用于 PCR 反应。

4.QIAGEN 纯化

QIAGEN 是以硅胶为基础的旋转柱。取 5 倍体积 QIAquick™缓冲液 PB 加到 DNA 样品中，混匀，将混合液加到 QIAGEN 旋转柱中，12 800r/min 离心 1min，去掉流出液，加 750μl QIAquick™缓冲液 PE 到旋转柱中，室温放置 1min，离心 1min，去流出液。加 100μl TE 到旋转柱中，离心 1min，收集流出液（100μl），取 5μl 用于 PCR 反应。

5.固相吸附法提取 DNA（Boom 法）

Boom 法原理是利用硫氰酸胍盐（GuSCN）溶解细胞和灭活核酸酶，DNA 被硅藻颗粒吸附，然后在低盐下将 DNA 从硅藻中解吸附出来，洗脱液可以直接用于 PCR 反应。Boom 法主要用于临床标本（血液、脓液、液化痰等）小量 DNA 的提取。

（1）试剂

硅藻悬浮液：取 10g 分析纯硅藻悬浮于 50ml 无菌超纯水中，再加入 32%（W/V）盐酸 500μl，高压灭菌后保存备用。

L6 裂解液：50mmol/L Tris－HCl，500mmol/L GuSCN，20mmol/L EDTA，0.1% Triton X－100

L2 洗涤液：120g GuSCN，100ml 0.1mol/L Tris－HCl（pH6.4）

70%乙醇、丙酮

（2）方法

取 100μl 液体标本，加 900μl L6 裂解液，20μl 硅藻液，旋涡振荡混合 10min，12 000r/min 离心 15s，去上清。在沉淀中加 1ml L2 洗涤液，12 000r/min 离心 15s，去上清。再重复洗一次。沉淀用 1ml 70%乙醇洗一次，再用 1ml 丙酮洗一次，去上清，56℃放置 10min 使丙酮挥发干净。加 60μl TE 缓冲液，56℃放置 10min，12 000r/min 离心 2min，收集 40μl 上清液，再加 40μl TE 洗硅藻－DNA 复合物，12 000r/min 离心 2min，收集上清液，与第一次的合并。此液可以直接用于 PCR 反应。

6. Chelex－100 结合硅藻提取方法

硅藻对 DNA 具有较好的吸附力，Chelex－100 提取物经过硅藻纯化，可除去许多 PCR 抑制剂和短的 DNA 片段，有效地提高扩增成功率。50μl Chelex－100 提取液或 PK 裂解液，加 900μl L6 裂解液（40mmol/L Tris－HCl，pH6.4，17.5mmol/L EDTA，1% TritonX－100，40mmol/L 硫氰酸〈thiocyanate〉），再加 50μl 20%硅藻（硅藻溶于 1%HCl），室温放置 5min，然后将溶液用注射器加在小柱上（Promega 公司），加 2ml 洗涤液（2.5mmol/L EDTA，pH8.0，5mmol/L Tris－HCl〈pH7.5〉，100mmol/L NaCl），60%乙醇洗一次，然后用 2ml 80%丙酮洗。将小柱转移到 1.5ml 离心管中，1 200r/min 离心 20s 以去除乙醇，室温下放置 10min，除净乙醇。在小柱中加 50μl 蒸馏水，并将小柱转移到另一干净离心管中，室温放置 5min 以释放 DNA，12 000r/min 离心 20s，收集流出的 DNA 溶液，溶液可以直接用于 PCR 反应或 4℃保存。

硅藻的提取率较低，1mm^2 血痕用 Chelex－100 提取可得到 250ng DNA，用 PK－硅藻和 Chelex－100－硅藻提取则为 15ng。因为硅藻对核苷酸和寡核苷酸结合力低，使总产量低。但硅藻能有效地除去 PCR 抑制剂，扩增重复性好。表 4－1 和表 4－2 是 Chelex－100 与 Chelex－100 加硅藻法的比较。

表 4－1 Chelex－100 与 Chelex－100/硅藻法提取 DNA 在 FES/FPS 基因座的扩增结果比较

载体	Chelex－100 提取液	PK 裂解液加硅藻处理	Chelex－100 提取液加硅藻
烟头	±	+	+
纸	+	+	+
木材	±	+	+
毛发	+	+	+
树皮	−	±	±
尼龙	+	+	+
绿色软纸巾	+	+	+
白色软纸巾	±	+	+
白色棉花	+	+	+
有色棉花	±	+	±
黑色斜纹布	−	±	±
绿色棉布	+	+	+
黑色棉布	±	+	+
仿皮革	−	+	+
蓝色羊毛	+	±	+
黑色丝绸	+	+	+

表 4-2 提取体积对 Chelex-100 与 Chelex-100/硅藻法提取 DNA 的扩增结果影响比较

载体	Chelex-100			Chelex-100/硅藻法		
	1μl	2μl	5μl	1μl	2μl	5μl
烟头	-	-	-	+	+	+
纸	-	+	+	+	+	+
木材	-	+	-	+	+	+
毛发	-	+	+	+	+	+
树皮	-	-	-	+	+	-
尼龙	-	+	-	+	+	+
绿色软纸巾	-	+	-	+	+	+
白色软纸巾	-	+	-	+	+	+
白色棉花	-	+	-	+	+	+
有色棉花	-	-	-	+	+	+
黑色斜纹布	-	-	-	+	+	+
绿色棉布	-	+	-	+	+	+
黑色棉布	-	-	-	+	+	+
仿皮革	-	-	-	+	+	+
蓝色羊毛	-	+	-	+	+	+
黑色丝绸	-	+	-	+	+	+

+表示得到较好扩增，-表示没有得到扩增

7. 硅珠 DNA 提取法

硅珠 DNA 提取法原理与固相法类似，在高浓度的硫氰酸胍的存在下，DNA 被二氧化硅特异性地吸附；被吸附的 DNA 在没有硫氰酸胍存在下，DNA 又从二氧化硅中释放出来。硫氰酸胍是高性能的蛋白变性剂，可以使蛋白质与 DNA 分离，在硅珠吸附前高速离心可以将变性的蛋白质等不溶物除去，吸附后的漂洗可将溶液中的 PCR 抑制物除去。因此提取的 DNA 比较纯，回收率高，可以从 0.5μl 的血液中提取到足以进行 PCR 反应的 DNA。

（1）试剂

吸附剂：12g GuSCN 溶于 10ml 0.1mol/L Tris-HCl（pH6.4），0.8ml 0.5mol/L EDTA，0.5ml Triton X100

漂洗剂：12g GuSCN 溶于 10ml 0.1mol/L Tris-HCl（pH6.4）

硅珠悬浮液：0.06g Silica 溶于 1ml 无菌超纯水中。

TES：1ml 1mol/L Tris（pH8.0），2ml 5mol/L NaCl，2ml 0.5mol/L（pH8.0）EDTA，95ml 无菌超纯水。

（2）方法

将 1～4μl 血液加入含有 100μl TES 和 20μl 十二烷基肌氨酸钠的 0.5ml 离心管中，煮沸 5min，旋涡混合 10s；再加入 300μl 吸附液，旋涡混合 5s，再加入 20μl 硅珠悬浮液，旋涡

混合5s，室温放置10～15min，旋涡混合5s，12 000r/min离心30s，去上清液。沉淀物加入200μl漂洗液，充分悬浮沉淀物，12 000r/min离心30s，去上清液，重复漂洗一次。沉淀物用70%乙醇漂洗2次，去乙醇，56℃放置10min；加入60μl TES，旋涡混合30s，56℃保温10min，12 000r/min离心10min，取1～20μl上清液直接进行PCR反应。

8. 碱裂解法提取DNA

利用0.2mol/L NaOH裂解细胞，然后用0.04mol/L Tris－HCl（pH7.5）中和，离心，上清液可以直接用于PCR扩增反应。不同种类的细胞，其0.2mol/L NaOH处理时间和温度有所不同，详细的提取方法分述如下：

（1）全血 取5μl抗凝血与500μl无菌蒸馏水混合，12 000r/min离心1min，收集沉淀，在沉淀中加入20μl 0.2mol/L NaOH，室温下放置5min提取DNA，然后加180μl 0.04mol/L Tris－HCl（pH7.5）中和，提取液可以直接用于PCR反应。

（2）血痕 取相当于含有5μl抗凝血的血痕，加入20μl 0.2mol/L NaOH，保证所有载体被碱溶液浸泡，75℃下放置5min提取DNA，然后加180μl 0.04mol/L Tris－HCl（pH7.5）中和，提取液可以直接用于PCR反应。

（3）精液 向1μl精液中加入20μl 0.2mol/L NaOH，再加入一滴石蜡油，以防止蒸发，75℃下保温30min提取DNA，然后加180μl 0.04mol/L Tris－HCl（pH7.5）中和，提取液可以直接用于PCR反应。

（4）精斑 取相当于含有1μl精液的精斑，加入20μl 0.2mol/L NaOH，保证所有载体被碱溶液浸泡，75℃下放置5min提取DNA，然后加180μl 0.04mol/L Tris－HCl（pH7.5）中和，提取液可以直接用于PCR反应。

（5）口腔表皮细胞 在室温空气中晾干口腔拭子，加入20μl 0.2mol/L NaOH，保证所有载体被碱溶液浸泡，75℃下放置10min提取DNA，然后加180μl 0.04mol/L Tris－HCl（pH7.5）中和，提取液可以直接用于PCR反应。

4.1.6 DNA提取方法的评估

法医DNA提取方法的好坏评估指标：一是DNA产量，二是DNA的质量，三是DNA的纯度。此外操作的繁复程度。实际应用是否可行及费用也都是被考虑的因素。DNA的质量常常比DNA产量更重要。提取的DNA适宜于下一步分析并获得良好的结果是选择提取方法最重要的标准。

不同的提取法，所获得DNA的量和质不同，费用也各不相同，表4－3和表4－4为各种方法的比较。Cattaneo（1997）对玻璃奶法（此方法将在线粒体DNA提取部分介绍）、乙酸钠和磁珠法提取陈旧DNA进行了比较，结果发现：提取的DNA扩增成功率最高的为玻璃奶法（65%），其次是乙酸钠（55%），最后为磁珠法（37%）。而提取的DNA量是乙酸钠（256μg/ml）比玻璃奶法（86μg/ml）多。玻璃奶法扩增效果好的原因可能是其能有效地去除PCR抑制剂。

表 4-3　不同 DNA 提取方法的 DNA 产率的比较

提取方法	方法原理	平均产率(ng)	范围	平均产率（ng）	范围
有机	有机/Centricon	1352	100～2100	730	280～1140
Chelex	离子交换树脂	2280	1200～3200	1840	800～2400
Definitive	蛋白酶	ND	ND	ND	ND
Dynabeads	生物磁性	930	600～1200	105	38～190
QIAamp	硅胶膜	1150	500～1050	260	60～500
READY-Amp	离子交换树脂	725	400～1050	570	400～900
Wizard	吸附	90	15～300	270	30～600

注：DNA 的产率来自 25μl 血液或血痕的　　ND：不能确定

表 4-4　不同 DNA 提取方法的 DNA 扩增率、时间与费用的比较

提取方法	扩增成功率(全血)	扩增效果(全血)	扩增成功率(血痕)a	扩增效果(血痕)b	整个时间 c(h)	净工作时间 d(h)	转移步骤数 e	每份提取费用(USA)
有机	3/3	+	4/4	+	4.75(5.5)	1.6(1.5)	3	4.5
Chelex	4/4	+	4/4	+	2.5	1.6	0	0.02
Definitive	0/4	-	0/4	-	0.75	0.6	0	1.5
Dynabeads	4/4	+	2/4	±	2.0(2.5)	1.9(2.0)	1	1.9
QIAamp	3/4	±	3/4	±	1.5(2.0)	1.4(1.6)	2(3)	3.89(4.5)
READY-Amp	4/4	++	4/4	++	2.0	1.6	0	2.5
Wizard	3/4	±	0/3	-	4.5	3.3	1	2.5

a）扩增成功数/分析数

b）扩增效果：-没有带，±较弱，+强，++很强

c）总时间指的是每一提取程序所需的总时间

d）净工作时间指的是提取过程中实际操作所花的时间

e）不包括取血步骤，（）内的为血痕的时间与步骤

在实际案件检验中常用的 DNA 提取方法主要是有机提取法与 Chelex-100 提取法。下面各类检材提取主要介绍这两种方法的提取。

4.2 各种检材 DNA 提取

4.2.1 试剂

（1）2×蔗糖-Triton X-100 细胞裂解液：0.64mol/L 蔗糖，0.01mol/L $MgCl_2$，20mmol/L Tris-HCl（pH7.6），1% Triton X-100，4℃保存

（2）EDTA-Na_2 提取液：75mmol/L NaCl，24mmol/L EDTA-Na_2（pH8.0），4℃保存

（3）10%SDS，室温放置

（4）饱和酚与氯仿：市售苯酚 179℃重蒸，加 8-羟基喹啉至终浓度 0.1%，以等体积的 0.8mol/L Tris-HCl（pH8.0）平衡至上层水相 pH 达 8.0，4℃闭光保存、备用

（5）3mol/L NaCl，生理盐水（0.9%）

（6）TE 缓冲液：10mmol/L Tris-HCl（pH 8.0），1mmol/L EDTA

（7）无水及 70%乙醇

（8）20%及 5%Chelex－100

（9）蛋白酶 K（Proteinase K）：2mg/ml 或 10mg/ml

（10）TNE：10mmol/L Tris－HCl（pH8.0），10mmol/L EDTA，100mmol/L NaCl

（11）5×精子 DNA 提取液：50mmol/L Tris－HCl（pH8.0），50mmol/L EDTA（pH8.0），500mmol/L NaCl，10%SDS，195mmol/L DTT

（12）0.4mol/L DTT

4.2.2 有机提取法提取各类检材 DNA

1. 血液

（1）3ml EDTA－抗凝血，加入等体积裂解液，剧烈振荡，使细胞裂解，溶液呈透亮状态。

（2）3 000～5 000r/min 离心 15min，去掉上清液，收集沉淀，用 3ml 生理盐水洗净沉淀，3 000～5 000r/min 下离心 15min，收集沉淀，如果此时沉淀仍带有红色，再反复用生理盐水洗沉淀，直至沉淀呈无色。

（3）加入 1.5ml EDTA－Na_2 提取液，悬浮沉淀，并反复吹打，使沉淀分散开来。

（4）加入 1/10 体积 10%SDS，使其最终浓度为 1%，混匀。

（5）加入 1/20 体积 2mg/ml PK，（最终浓度 100μg/ml），混匀，封口，于 37℃水温保温 4h。

（6）向溶液加入等体积的饱和酚，以上下反复颠倒试管方式混匀液体，将管内溶液混合成乳状液。3 000～5 000r/min 下离心 5min，此时溶液分成三层，DNA 在无色的上层水层中，中间白色混浊层为变性蛋白质。如果 PK 水解蛋白质彻底，此层物质就会变得少些，最下层带褐或黄色的是酚层。

（7）转移上层水溶液到一干净管中。为增加 DNA 回收量，可以在有机相中加入等体积 TE 溶液（pH7.8），充分混匀后，离心，吸取上层水相，与第一次抽提的水相合并。

（8）加入 1/2 体积的酚和 1/2 体积的氯仿，混匀。3 000～5 000r/min 下离心 5min。如果水相层和有机层的界面不甚清楚，说明其中含蛋白质较高，可采用多次酚/氯仿溶液抽提，并可适当加大离心速度和或延长离心时间。此步中可以加少许异戊醇，目的是减少蛋白质在变性操作过程中产生气泡。

（9）转移上层水溶液到另一干净管中，加入等体积的氯仿混匀，3 000～5 000r/min 离心 5min。

（10）转移上层水溶液到另一干净管中，加入 1/10 体积的 3mol/L NaCl 轻轻混匀，加入 2.5 倍溶液体积的冰冷（－20℃保存的）无水乙醇，混匀，沉淀 DNA。

（11）10 000r/min 离心 5min，收集沉淀，用 70%乙醇洗去 DNA 中的盐，5 000r/min 离心 5min，除去液体，真空抽干 DNA 5min，至 DNA 成无色透明状（注：干燥时间随使用真空度变化，或长或短）。

（12）加适量（50～100μl）TE 缓冲液溶解 DNA，溶解后的 DNA 溶液于 4℃保存。

或在等体积酚－氯仿，氯仿抽提后，将上清液转移到 Centricon30（Amicon，Inc.，Beverly，MA）超滤管中，2 500r/min 离心 25min，用 2ml 无菌水冲洗截留液，2 500g 离心 25min，重复 3 次，以除去 DNA 溶液中的 SDS、盐和一些杂质，纯化 DNA。

2. 血斑

（1）将血斑剪碎，加入适量 TNE 浸泡底物，TNE 的体积以底物全部浸湿，底物表面还保留 1～2mm 左右高度的 TNE 溶液层为宜，加入 1/10 体积的 10%SDS 和 1/20 体积的 PK（2mg/ml）使终浓度分别为 1%和 100μg/ml，37℃保温 3 小时。

（2）用套管法离心（5 000r/min 离心 5min）除去载体底物。具体操作方法：将底部带小孔的小管套在无小孔的管上，将浸泡好的载体及溶液加入带孔的管中，离心分离溶液与载体物。也可以不用除去载体，直接加酚抽提。

（3）收集溶液，用酚－氯仿抽提，以下步骤同 4.2.2.1 血液样品（6）～（11）。

皮革、木头或塑料制品上的血痕最好转移下来进行提取，以防止这些载体内可能存在的酶抑制物干扰内切酶或 Taq DNA 聚合酶活性。

3. 精液（精斑）

（1）取 0.5ml 精液加入等体积生理盐水，洗涤 2～3 次，3 000r/min 离心 5min，收集沉淀物，加入精子 DNA 提取液和蛋白酶 K（终浓度为 100ug/ml）混匀，于 37℃水浴保温消化 3h。DNA 提取及溶解过程同 4.2.2.1 血液样品（6）～（11）。

（2）从精斑中提取 DNA

方法同 4.2.2.2 血斑样品。

4. 从精液和阴道分泌液混合斑中提取精子 DNA

（1）精子检查：取少许检材经伊红美蓝染色，于显微镜下观察，确认精子的存在。

（2）将带有精子的检材剪碎，装入 1.5ml 离心管中，加 1ml TNE 于 37℃浸泡 30min。然后加入 1/10 体积的 10% SDS 至终浓度为 1%，加蛋白酶 K 至终浓度 40μg/ml，于 37℃保温 2h 消化上皮细胞。对于那些需要对女性阴道上皮细胞进行鉴定的物证，加 TNE 后先在 37℃保温 1～2h，通过离心收集细胞，使女性上皮细胞和精子等细胞先与载体分离开来，此时不要加 SDS 和 PK。然后在收集的细胞中加 SDS 和 PK 进行分步提取。

（3）将消化后的混合物按血斑中的套管法离心去载体，3 000～5 000r/min 离心 15min，收集流出液中的沉淀物。

（4）用 TNE 洗沉淀物，3 000～5 000r/min 离心 15min，重复此步骤一次，收集沉淀物。

（5）向沉淀物中加入精子提取液和蛋白酶 K，终体积为 300μl，PK 终浓度为 100μg/ml，于 37℃水浴中消化 3h。

（6）DNA 抽提步骤同 4.2.2.1 血液样品 DNA 提取（6）～（11）。

消化上皮细胞的时间与 PK 用量根据检材所含上皮细胞与精子的情况而定。如果女性物质较多，PK 可增至 200μg/ml，延长消化时间，或进行二次消化。浸泡和消化用的 TNE 也可以由蒸馏水或磷酸（PBS）缓冲液代替。

5. 人体有核细胞组织

取组织约 0.1g，用生理盐水洗净附着的血液后将检材剪碎，加入 DNA 提取缓冲液，用匀浆器研磨，再加入蛋白酶 K（终浓度 200μg/ml）和 1/10 体积的 10% SDS，于 37℃消化 3h。随后 DNA 提取步骤同 4.2.2.1 血液样品（6）～（11）。

6. 唾液（斑）

同血液（斑）。

7. 骨骼、牙齿

骨骼、牙齿由于有外部硬组织的保护，因此能够抵抗外界环境的影响破坏，使其中的DNA不易被降解，在机体其他部分均腐败或炭化的情况下，仍可以从骨骼、牙齿中提取到DNA，进行DNA分析。已有报道从数千年前的白骨中提取到DNA，成功地进行了DNA分析。骨髓腔中的骨髓、骨密质及骨松质均含有DNA，其中骨髓中DNA含量最丰富、骨松质DNA含量为骨密质10～20倍。每克骨松质含1.5～10μg DNA，每克骨密质含0～0.5μg DNA。但不同保存时间、保存条件，骨骼DNA含量有所不同。因此硬组织DNA提取的顺序一般是首选骨髓、牙髓，其次选择骨干骺端、其他骨松质丰富的部位，最后选择骨密质部分。骨骼DNA提取常采用脱钙与非脱钙两种方法，脱钙与否，虽不影响PCR扩增结果，但不脱钙骨的DNA提取量是脱钙的两倍。

通常现场提取到的骨骼、牙齿标本都是长期处于外界环境中的，在发现、采集及送检过程中，可能受到一些污染，得到这类标本首先要进行清洗。具体做法是将标本用水冲洗，手术刀刮净骨块或牙齿表面物，再用水洗干净、干燥，紫外光（254nm）照射1.5h。此步旨在除去检材外表的各种污染物，以及在物证的发现和提取过程中的污染。

（1）骨髓 DNA 提取

新鲜骨头，采集骨髓0.5～1g，液氮下冷冻，研碎，加入1ml TNE缓冲液，1/10体积的10%SDS和1/20体积的PK（终浓渡100μg/ml）混匀，56℃保温过夜，以后同4.2.2.1血液样品（6）～（11）步骤。

（2）骨松质样品 DNA 提取

以手术刀刮净骨骼和牙齿表面垢物后，再用温热水冲洗两次，紫外光照射1h。然后用锤子将牙齿砸成粉末，或用电钻取出骨松质，并将其用研磨器进一步磨碎。

方法1：

取骨或牙齿粉末约2g，放入5ml的试管中，加入DNA提取缓冲液（10mmol/L Tris－HCl，pH8.0，10mmol/L EDTA，100mmol/L NaCl，2% SDS，39mmol/L DTT）充分混匀，再加入200μg/ml蛋白酶K，于37℃保温过夜，次日用酚和氯仿提取，将提取液置Centricon 100中，2 500r/min离心15min，留下浓缩液后，加1/10体积的3mol/L NaCl和2.5倍无水乙醇沉淀并置－20℃冰箱过夜。次日于4℃下13 000r/min离心20min、干燥，加40μl TE溶解。最后DNA溶解液再经DNA clean－up纯化系统。

clean－up 纯化 DNA：

（1）在DNA溶解液中加入1ml Resin（试剂盒内试剂）混匀后加入到安装有纯化柱的5ml注射器，将液体缓慢推出。

（2）加入2ml 80%异丙醇，将液体缓慢推出。

（3）取下柱子于13 000r/min离心20s，置柱子室温5～10min，向柱内加入30μl TE，静置5min，于13 000r/min离心20s，得到纯化的DNA样品，4℃备用。

方法2：

（1）取骨粉末约2g放入研钵中，加入2ml的CTAB提取缓冲液（2%CTAB，20mmol/L Tris－HCl，pH8.0，1mmol/L EDTA，1.4mmol/L NaCl，0.2%巯基乙醇），不断研磨直到成稀糊状，用刀片刮取到5ml的试管中，再加入1ml的CTAB提取缓冲液于旋涡振荡器上充

分混匀，置室温过夜。

(2) 次日将上述样品在旋涡振荡器上充分混匀2min，置65℃水浴1h，其间不断取出在旋涡振荡器上充分混匀。

(3) 将样品在10 000r/min离心25min，在上清液中加入等体积的氯仿/异戊醇（24∶1），充分混匀，于10 000r/min离心10min，保留上清液。如上清液超过2ml体积进行下步浓缩，反之则不需要。

(4) 将提取液置Centricon 100中，500r/min离心15min，留下上管的浓缩液。重复此步骤，直到提取液全部浓缩完。

(5) 将浓缩液与QIAquick™ PCR纯化试剂盒中的5倍体积的PB溶液混匀，直接放在试剂盒提供的纯化柱上，于台式离心机中13 000r/min离心1min，去掉下层液体。重复此步骤直到液体全部过完柱。

(6) 加750μl PE溶液（试剂盒），于台式离心机中13 000r/min离心1min，去掉下层液体。再以同样的方式离心一次。在室温挥发5min。

(7) 将柱子放在一干净的1.5ml试管上，取30μl或50μl EB溶液（试剂盒）加到柱子上，70℃放置10min，于台式离心机中13 000r/min离心1min，收集下层液。

(8) 收集的下层液置于Centricon 100中进一步浓缩，方法同上。浓缩液4℃保存。

对于条件较好的骨骼加30μl的EB溶液一次即可，也不需再浓缩。对于条件不好的骨骼，先加50μl的EB溶液到柱子上，70℃放置10min，于台式离心机中13 000r/min离心1min，收集下层液。再用50μl的EB溶液到柱子上，70℃放置10min，于台式离心机中13 000r/min离心1min，收集下层液。两次的液体合并，用microcon100或Centricon 30进一步浓缩。

(3) 牙髓样品DNA

牙齿的牙髓腔内含有牙髓，有血细胞和淋巴细胞等，是提取牙齿DNA的最好部位。

牙齿用水洗净，用干净无菌锯将牙纵面锯开，或用干净纸包好牙齿，用锤子砸牙齿，使牙齿裂成两半。取出牙髓置子小试管内，用无菌不锈钢小药匙轻压牙髓，加入500μl TNE（0.01mol/L Tris-HCl pH7.5，0.1mol/L NaCl，0.01mol/L EDTA），SDS其终浓度为2%，加入PK使终浓度为100μg/ml，56℃保温过夜，以下同4.2.2.1血液样品（6）~（11）步骤。

(4) 象牙质中DNA

①用牙科工具将牙纵向断裂，切成厚1mm的切片，再度切割，采牙髓及象牙质部分。

②象牙质用无菌水超声波10min，洗净后用0.5mol/L EDTA于室温下脱钙2~7天。

③脱钙后的象牙质加入1ml TNE加SDS和PK使最终浓度分别为2%和100μg/ml，慢慢搅拌均匀，56℃保温过液，以下同4.2.2.1血液样品（6）~（11）步骤。

对于较为陈旧的骨骼，或将标本研碎成粉末，或用牙科钻将骨钻成粉。称取300mg，溶于1.5ml 0.5mol/L EDTA（pH8.3），室温下振荡48h脱钙，以较长时间保持均匀混合状态以便充分吸收骨粉中的无机盐，使包围在其中的有机成分稀释出来，12 000r/min离心取上清液。脱钙也可以在50℃脱钙12h。

骨、牙等硬组织样品DNA提取注意点：①留意样品腐败状态，在骨髓腐败情况下，

不可能提到骨髓 DNA，此时，应利用骨松质。一般在肋骨，大腿骨、上腕骨等骨端常残留骨髓。②骨松质 DNA 随死后时间加长而减少，一般干燥状态下保存时间较长的白骨也能得到 DNA，水中或土中放置 1～3 个月时间也能得到 DNA，但时间再长，很难得到 DNA。腐败牙齿、烧死尸体的牙齿（由于加热），几乎提不到 DNA。

骨髓或牙髓 DNA 提取的也可以用 Chelex－100 方法，详见 Chelex－100 提取牙髓和骨髓部分。

8. 甲醛固定组织

甲醛（福尔马林）是法医或医院保存组织标本的常用防腐剂，它深入细胞，使细胞膜蛋白变性，并使 DNA 与核蛋白结合。从甲醛固定组织提取 DNA，关键是破坏已变性的细胞膜蛋白，打开 DNA 与核蛋白之间的连接，使 DNA 恢复自由状态。研究发现，适当提高温度，有利于打开 DNA 与核蛋白之间的连接。

提取操作：剪取甲醛固定的组织如肌肉，用蒸馏水冲洗 2～3 次，以无菌剪刀或刀片弄碎或冰水浴匀浆，用 2×蔗糖－Triton X－100 裂解液（0.6mol/L 蔗糖，0.02mol/L Tris－HCl，pH8.0，0.01mol/L $MgCl_2$，2% Triton X－100）清洗，3 000r/min 下离心 15min，收集沉淀。在沉淀中加入 1.5ml TNE（0.5mol/L Tris－HCl，pH8.0，0.02mol/L EDTA，0.01mol/L NaCl），混匀，加入 10% SDS 和蛋白酶 K，使终浓度分别为 1% 和 200μg/ml，48℃水浴保温 48h，如果观察溶液仍有许多不溶物，可增加 SDS 和 PK 或延长时间。以下步骤同 4.2.2.1 血液（6）～（11）步。

甲醛固定组织 DNA 的提取，关键是要把甲醛去除干净。甲醛使 DNA 变性，因此甲醛固定的组织很难得到高分子 DNA。

9. 石蜡包埋组织

（1）将微量石蜡包埋组织放入 1.5ml 的 Eppendorf 管中，加入 1ml STE 液（10mmol/L Tris－HCl，pH7.5，100mmol/L NaCl，1mmol/L EDTA），置 72℃水浴 10min，插入吸水滤纸条，将浮蜡及 STE 吸干，再加入 1ml STE，反复多次，至石蜡除净。或用二甲苯脱蜡，具体方法将在 4.2.3.12 提取部分介绍。

（2）将组织研磨成悬液，加 EDTA－Na_2 提取液 1ml，加 SDS 和蛋白酶 K 至终浓度 1% 和 100μg/ml，置 37℃水浴保温 4h 或过夜。

（3）以后酚－氯仿抽提同 4.2.2.1 血液样品（6）～（11）。

10. 鼻涕

鼻涕样品 DNA 提取同 4.2.2.2 血斑样品方法，只是将消化液中的 SDS 改为月桂酰钠，其他相同。

4.2.3 微量检材 DNA 提取

1. 血液

（1）取 4～10μl EDTA 抗凝血，加到 0.5ml 离心管中，再加入 500μl 灭菌蒸馏水，剧烈振荡，室温下放置 15min。

（2）13 000r/min 下离心 3min，去上清液，收集沉淀，反复用无菌蒸馏水洗净白细胞沉淀物，直至无色或血色素很少。

（3）加入 200μl 5% Chelex－100 溶液（5% Chelex 为悬浊液，使用前要振摇悬浮 Chelex

颗粒，在旋涡振荡器上振荡7秒钟，56℃保温30min或更长，振荡7s，100℃保温8min后，剧烈振荡7s，13 000r/min下离心3min，以沉淀Chelex颗粒，上清液可用于PCR反应或者4℃保存。

2. 血斑

（1）取0.5cm^{2}的血痕加到小离心管中，加入500μl无菌蒸馏水，剧烈振荡，室温下放置15min。

（2）13 000r/min下离心3min，去上清液，收集沉淀，反复用无菌蒸馏水洗净白细胞沉淀物，直至无色或血色素很少。载体不需要去除，始终留在离心管内。

（3）以下步骤同4.2.3.1血液（3）。

3. 混合精斑

按差异裂解分离法处理除去混合斑中上皮细胞，干净的精子沉淀物用200μl 5%Chelex溶液悬浮，加DTT和PK至终浓度分别为0.039mol/L和100μg/ml，在旋涡振荡器上振荡7秒钟，56℃保温过夜，振荡7秒钟，100℃保温8min后，剧烈振荡7秒钟，13 000r/min下离心3min，以沉淀Chelex颗粒，上清液用于PCR反应或者4℃保存。

4. 唾液（斑）

唾液（斑）样品的Chelex法处理同4.2.3.2血斑。

5. 烟头、口香糖和其他含有检材的唾液斑

（1）烟头 以干净无菌镊子夹住烟头，用干净无菌的剪刀剪下烟头外圈的纸，并剪碎放入0.5ml的离心管中，加入500μl无菌水洗烟头纸，13000r/min离心3min，去上清，在收集的沉淀中加入200μl 5%Chelex-100悬浮液，56℃保温30min或更长时间，100℃保温8min，剧烈振荡7s，13 000r/min离心3min，上清液可以直接用于PCR分析。

（2）口香糖 将口香糖剪碎，放入0.5ml的离心管中，加入500μl无菌水洗口香糖，13 000r/min离心3min，去上清，在收集的沉淀物中加入200μl 5% Chelex-100悬浮液，56℃保温30min或更长时间，100℃保温8min，剧烈振荡7s，13 000r/min离心3min，上清液可直接用于PCR分析。

（3）其他含有唾液斑的检材如邮票、信封等DNA提取同4.2.3.5（1）烟头。

6. 毛根

剪取毛发的毛根部分5~10mm，置于1.5ml离心管中，加入200μl 5% Chelex-100和2μl 10mg/ml蛋白酶K，56℃保温过夜（至少6hr），剧烈振荡5~10s，沸水煮8min，再剧烈振荡5~10s，13 000r/min离心2~3min，取20μl上清液用于扩增反应。

7. 其他有核组织

取冷冻组织少许，研碎，用生理盐水洗去附着在组织上的血，5 000r/min离心5min，去上清液；用200μl 5%Chelex-100悬浮沉淀，加2μl 10mg/ml蛋白酶K，56℃保温过夜（至少3hr），剧烈振荡5~10s，沸水煮8min，再剧烈振荡5~10s，13 000r/min离心2~3min，取1μl上清液用于PCR反应。

8. 尿液

尿液中含有一些脱落的上皮细胞，男性尿液中有时会含有一些精子，但含量较少，DNA含量低，一般男性尿液DNA量为4~60ng/ml，女性为14~200ng/ml。尿液成分复杂，

含有一些DNA降解物，室温保存7天则无法进行DNA检验，在尿液中加EDTA可以提高DNA的回收率。

提取操作：取尿液1～5ml，1 500r/min离心15min，收集沉渣，去上清。沉淀中加0.3ml NaCl，悬浮，13 000r/min离心5min，收集沉渣。然后加200μl 5% Chelex－100悬浮液，56℃保温30min或更长时间，100℃保温8min，剧烈振荡7s，13 000r/min离心，上清液可用于PCR分析。

9. 羊水

取100μl羊水，15 000r/min离心3min，弃上清，加入200μl 10% Chelex－100和2μl 10mg/ml蛋白酶K，混匀；56℃保温30min；高速振荡5～10s，15 000r/min离心10～20s，沸水煮8min，高速振荡5～10s，10 000～15 000r/min离心2～3min，取上清液用于PCR反应。

或用碱裂解法提取羊水DNA：取羊水1.5ml，15 000r/min离心3min，弃上清，沉淀用TE缓冲液（10mmol/L Tris－HCl，1mmol/L EDTA，pH7.5）洗涤一次后用50μl裂解液（0.1mol/L NaOH，2mol/L NaCl）悬浮沉淀物，100℃保温2min，15 000r/min离心3min，上清用于PCR反应。

10. 骨髓和牙髓

按4.2.2中骨骼方法处理骨骼和牙齿，取出骨髓或牙髓，研碎的牙髓或骨髓，加入200μl 5% Chelex－100和PK，使PK终浓度为100μg/ml，56℃保温过夜，振荡7s，100℃保温8min，振荡7s，13 000r/min离心3min，上清液用于PCR反应。

11. 组织切片

（1）加热融化封片的树胶，揭去盖玻片。

（2）用手术刀片将切片组织及树胶一同刮下，移入1.5ml离心管内。在离心管内加入200μl二甲苯，振荡混匀，10 000r/min离心5min，弃上清。在沉淀中加入200μl无水乙醇，振荡混匀，10 000r/min离心5min，弃上清。重复本步骤1次。

（3）在沉淀中加入1滴丙酮，开盖，置56℃温浴中，使液体完全挥发。

（4）在处理后的切片组织中加入100μl消化液（50mmol/L Tris－HCl，10mmol/L NaCl，10mmol/L EDTA，0.5% Tween－20）及2μl 10mg/ml PK，2.5μl 20% SDS，轻轻混匀，56℃温浴3h后，转入37℃温浴12h以上，95℃温浴10min，10 000r/min离心5min。上清用于PCR反应。

12. 石蜡包埋组织

用手术刀片将组织切下，移入1.5ml离心管内。在离心管内加入1ml二甲苯，室温脱蜡1h，重复一次，10 000r/min离心5min，弃上清。在沉淀中加入1ml无水乙醇，振荡混匀，10 000r/min离心5min，弃上清。重复本步骤1次，室温下晾干。也可以不离心，直接将蜡块取出。

将处理后的组织切碎，加入100μl Chelex－100悬浮液及2μl 10mg/ml PK，56℃温浴3h后，100℃温浴8min，10 000r/min离心5min。上清用于PCR反应。

13. 头皮屑

（1）取深色洁净纸一张，将附着有头皮屑的物品置于其上，用小毛刷轻刷。

(2) 将收集的头屑放入 0.5ml 离心管内，加入 10μl 碱裂解液（200mmol/L KOH，50mmol/L DTT)，65℃温浴 30min，加入 10μl 中和液（900mmol/L Tris - HCl，pH8.3，200mmol/L HCl)，混匀离心。上清用于 PCR 反应。

头皮屑中有核细胞数量受生理状况影响较大，因此有些头屑中能够得到 DNA 分型结果，有些则不能。

14. 汗潜指纹及已显现指纹

指纹由于其具有个人认定作用，在侦察破案中一直受到重视而成为现场重点被搜集的物证。在遗留指纹的细胞成分中同样存在有 DNA，可进行 DNA 分析。

(1) 用无菌水将小块脱脂棉稍微沾湿后涂擦已显现的指纹。

(2) 将涂擦的棉球放入 1.5ml 离心管内，用 1ml 无菌水浸泡，室温放置并不断搅动。待指纹显现试剂从棉花上充分溶解、脱落后，取出棉花。13 000r/min 离心 3min，弃上清；沉渣用无菌水清洗，重复离心步骤 3 次。沉渣用 200μl 5% Chelex - 100 充分悬浮，56℃温浴 30min，取出振荡，100℃温浴 8min，充分振荡。13 000r/min 离心 3min，上清用 Microcon - 100 柱过滤浓缩。浓缩液用于 PCR 扩增。

15. 毛干线粒体 DNA

将毛干剪成约 3mm 长，放入 0.5ml 离心管中，用无水乙醇冲洗 1 次，用水冲洗 1 次，用无水乙醇冲洗 1 次，晾干后加入 0.4mol/L DTT 溶液 3μl，10%SDS 溶液 3μl，2mg/ml 蛋白酶 K 溶液 6μl，EDTA - Na_2 提取缓冲液 18μl，混匀后 37℃水浴保温待毛干溶化，大约需 5h。然后可以使用以下三种方法之一进行 DNA 提取处理。

(1) 有机溶剂法：

将毛干消化液以等体积酚抽提一次，再以等体积的 1:1 混合的酚:氯仿抽提一次，最后以等体积的氯仿抽提一次，然后加入 1/10 体积 3mol/L 乙酸钠和 2.5 倍无水乙醇，混合后 - 20 ℃冷冻 1h，然后在台式高速离心机上 16 000r/min 离心 20min，弃上清后用 70% 乙醇冲洗 1 次，凉干后加入 10μl TE 缓冲液备用。

(2) 玻璃奶法：

将毛干消化液中加入 3 倍体积的溶胶液后混匀，加入 10μl 玻璃奶，混匀后室温放置 10min，然后以 7 000r/min 离心数秒钟，弃上清液后加入 125μl 洗脱液，混匀后，7 000r/min 离心数秒钟，重复两次洗脱，然后加入 20μl TE 缓冲液，混匀后 60℃保温 10min，15 000g离心数秒钟，收集上清液备用。

(3) Chelex - 100 法

将毛干消化液中加入 200μl 20% Chelex - 100 液，混匀振荡 1min 后，煮沸 8min，再混匀振荡 1min 后，12 000r/min 离心 3mim。或直接将洗干净的毛发放入含 200μl 5%Chelex - 100 溶液中，加 0.4mol/L DTT 和 PK 使终浓度分别为 0.04mol/L 和 100μg/ml，56℃保温过夜，振荡数秒，100℃保温 8min，振荡数秒，13 000r/min 离心 3min，上清液可用于 PCR 反应。

(4) PCR 缓冲液处理法

它是目前较常用的方法。具体方法：洗干净的毛干加 50μl 1 × PCR 反应缓冲液，0.4mol/L DTT 溶液 5μl，2mg/ml PK 溶液 5μl，混匀后 56℃水浴保温待毛干溶化，13 000r/min 离心 3min。取上清液 5μl 用于 PCR 反应。

16. 指甲线粒体 DNA

指甲中细胞已角化，很难提取到核 DNA，但存在有线粒体 DNA。

提取操作：指甲在蒸馏水中先煮 5min，剪成小片，用刀片刮削指甲成碎屑片可以加快消化，将指甲碎片放入 0.5ml 离心管中，加 0.5ml 消化液（10mmol/L EDTA，50mmol/L NaCl，2%SDS（w/v）），再加 20μl 1mol/L DTT，15μl 20mg/ml PK，56℃保温 6h 或过夜，再加新鲜 20μl 1mol/L DTT，15μl 20mg/ml PK，56℃保温 6h 或过夜。消化液用 Microcon 100 过滤，截留物用 TBE（1mol/L Tris，20mmol/L EDTA - Na_2）冲洗 3 次，取 40μl 截留物用于扩增。

或将指甲剪碎后，同 4.2.3.12 毛干方法处理提取 DNA。

17. 其他检材线粒体 DNA

各种组织体液细胞线粒体 DNA 提取方法除毛干、指甲外，其他检材均同核 DNA 的提取方法，只是对于血痕等检材用水少洗几次，因为线粒体存在于细胞质中。

18. 染色涂片

用无菌水湿润的干净无菌的纱线擦拭染色涂片，放入 0.5ml 离心管，加 400μl 无菌水，振荡使染料尽可能从纱线中脱落下来溶于水中，13 000r/min 离心 5min，收集沉淀。如果是精斑的染色涂片，则以下步骤按混合斑检材提取。如果是其他细胞的涂片，则可按白细胞处理方法提取 DNA。

19. 脱落细胞

按唾液斑的提取方法提取，如 Chelex - 100 或碱裂解法提取 DNA。

20. 大便中 DNA

用提取泥土中 DNA 的试剂盒——UltraClean soil DNA Isolation KitTM（由美国加州 Solana Beach 的 Mo Bio Laboratories 生产）提取大便中的 DNA。

4.2.4 提取注意事项

（1）酚抽提时如果上清液太粘，不能和蛋白质分开时，可以加入适量 TES（15mmol/L Tris pH8.0，15mmol/L EDTA pH8.0，15mmol/L NaCl）稀释，然后再用酚抽提。

（2）温浴可在 45～55℃范围内进行，亦可 37℃过夜。

（3）抽提 DNA 时动作不可过猛，以防机械震动将 DNA 分子打断。

（4）沉淀 DNA 时要小心操作，勿使絮状物断成小丝。

（5）沉淀 DNA 时，可不加 3mol/L 醋酸钠（pH5.2），但乙醇溶解部分小分子的 DNA，影响 DNA 回收率。

（6）70%乙醇加至离心管的 2/3 高度，为防止 DNA 损失可用旋涡振荡器稍加振荡，然后按上述条件离心回收 DNA。经 70%乙醇洗过的 DNA 沉淀不再粘在管壁上，因此去上清液需特别小心。

（7）DNA 溶于 TE 时可先浓一点，发现太浓时再加 TE。这样能保证 DNA 样品不致太稀而无法使用。一般来讲，浓度为 0.4～0.6μg/μl 时较为理想。

（8）若血痕极稀薄，可减少 Chelex - 100 体积，使其浓度相应提高以达到扩增所需模板量。对于较浓的血痕则可增加 Chelex - 100 体积或减少血痕用量使 PCR 反应体系内的模板量达到最佳。

(9) 在检验混合斑第一步消化离心后，涂片显微镜检查。若视野内存在一些上皮细胞碎片，没有完整的细胞核，属正常现象；若仍有完整的上皮细胞核则必须重新进行细胞差异裂解法进行分离。

(10) 在精子量极少的情况下，加入冷无水乙醇可能看不到絮状的 DNA，此时应将样品置于 -20℃过夜，10 000r/min 离心 10min，倒掉上清即可看到管底沙粒状的 DNA 沉淀。

(11) 匀浆，研磨虽然能破碎大块组织，以利于更好的去除蛋白质，但同时也会因匀浆这种机械操作使 DNA 受到损伤而发生断裂。因此应尽量温和操作，以减小对 DNA 分子的剪切力。

(12) 酚会造成皮肤的严重烧伤及衣物损坏，所以凡接触酚的时候必须戴手套、穿工作服。操作应在通风柜中进行。

(13) 实验废液，如酚、氯仿等必须用专门容器收集，统一处理。

4.3 DNA 浓缩

4.3.1 沉淀浓缩法

DNA 是多聚两性离子的水溶性化合物，它与钠、钾、镁形成的盐在许多有机溶剂中不溶解，但不会被有机溶剂变性，由此可从溶液中收集 DNA。常用的有机溶剂有乙醇、异丙醇、聚乙二醇等。DNA 沉淀是 DNA 浓缩最常用的方法，其最大的优点是通过 DNA 沉淀来改变 DNA 的溶解缓冲液及重新调节 DNA 溶液的浓度，也可去除溶液中某些盐离子与杂质，在一定程度上纯化 DNA。乙醇沉淀是浓缩 DNA 最常用的方法。使 DNA 在低温（-20℃）和中等浓度一价阳离子存在下形成沉淀后，离心收集，并按所需浓度重新溶于适当缓冲液。此法快捷而且沉淀完全，即使只有几 μgDNA 也能回收。

1. 沉淀用盐

在 DNA 沉淀中使用最多的是金属 1 价离子，如 Na^+、K^+ 和 NH_4^+，2 价 Mg^{2+} 的沉淀效率最高。每一种离子均有各自的优缺点和应用范围，并且在不同的有机溶剂中其使用的浓度也不一样（表 4-5）。

表 4-5 DNA 沉淀的盐浓度

盐	贮存液（mol/L）	终浓度（mol/L）
$MgCl_2$	1	0.01
NaAc	3.0（pH5.2）	0.3
KAc	3.0（pH5.2）	0.3
NH_4Ac	10.0	2.0~2.5
NaCl	5.0	0.2

醋酸钠　为沉淀 DNA、RNA 的最常用的盐类。终浓度为 0.3mol/L（pH5.2）。

氯化钠　对于含有 SDS（十二烷基硫酸钠）的 DNA 样品，最好选用 NaCl 沉淀，终浓度为 0.2mol/L，SDS 在 70%的乙醇中可保持溶解状态，不与 DNA 共沉淀，从而可通过弃上清去除这种去污剂，避免对以后酶促反应的影响。

醋酸铵（10.0mol/L） 四种三磷酸脱氧核糖核酸（dNTP）在醋酸铵盐溶液中，具有较大的溶解度，通过乙醇沉淀DNA，可去除大部分dNTP。有人证明，在2～2.5mol/L醋酸铵存在下，连续两次乙醇沉淀，可除去99%的dNTP，而DNA回收率在90%以上。

醋酸钾 沉淀效果同醋酸钠，但DNA的钾盐形式很难溶于含SDS的溶液，在下一步溶解DNA沉淀的缓冲液中含有SDS的情况下，则不能选用钾盐的方法沉淀DNA。

氯化镁 Mg^{2+}虽是核酸沉淀中最有效的离子，但氯化镁沉淀的DNA很难溶解，而且含Mg^{2+}的DNA溶液在保存过程中易降解，因为Mg^{2+}是DNA酶的激活剂。其实DNA沉淀后用70%乙醇漂洗两次，基本上可以去除痕迹量的Mg^{2+}，对于大量DNA，可先用适当体积的水溶解，然后补充Tris－HCl与EDTA至TE缓冲液浓度即可解决。

另外，DNA溶液在高于表4－5的盐浓度时，应该用TE（pH8.0）进行稀释，以避免盐与DNA一起沉淀。大于10mmol/L浓度的EDTA或磷酸盐肯定会在乙醇中沉淀，对于这两种情况，需要先透析或过柱去除盐离子，然后再沉淀。

2.DNA沉淀的温度与时间

DNA沉淀均应严格地在低温条件下进行，并推荐在－20℃或－70℃条件下放置数小时，甚至过夜。但目前有人认为，此点并无必要。因为低温限制DNA分子的布朗运动，所以在沉淀时低温条件并不是至关重要的因素。即使DNA浓度低到20ng/ml，长度短至20～30bp，在没有载体存在下，在0℃甚至室温下，10min即可形成沉淀。另外低温与长时间的沉淀，很容易导致盐与DNA共沉淀，影响下一步实验。一般情况的DNA沉淀，在0℃冰水放置10～15min足以达到实验的要求（表4－6）。

表4－6 温度与DNA的乙醇沉淀

DNA浓度	离子浓度	温度	沉淀效率（%）
0.1μg/ml	10mmol/L $MgCl_2$	0℃	35
	10mmol/L $MgCl_2$	室温	75
	1mol/L NaCl	0℃	18
10μg/ml	10mmol/L $MgCl_2$	0℃	95
	10mmol/L $MgCl_2$	室温	95
	1mol/L NaCl	0℃	95

3.有机沉淀剂

DNA沉淀中常用的有机溶剂有乙醇、异丙醇和聚乙二醇等。

乙醇 乙醇是首选沉淀DNA的有机溶剂。它对盐类沉淀少，DNA沉淀中所含的痕迹量乙醇易蒸发去除，不影响下一步实验。在适当的盐浓度下，2.5倍样品体积的95%乙醇可有效沉淀DNA。

异丙醇 其优点是所需体积小且速度快，适用于低浓度而体积大的DNA样品沉淀。0.54～1.0倍的异丙醇可选择性沉淀DNA、大分子rRNA和mRNA。但对5sRNA、tRNA及多糖不产生沉淀，一般不需要在低温条件下长时间放置。其缺点是易使盐（如NaCl）及蔗糖与DNA共沉淀，尤其是在－70℃沉淀时；DNA沉淀中存在的异丙醇挥发性差，较难

除去残留的异丙醇。所以常规需要用70%乙醇漂洗DNA沉淀物数次。如果没有必要减少体积，最好用乙醇沉淀。

此方法适合小体积或大体积DNA的提取，提取的DNA纯度高，量多、大分子多，可用于RFLP和PCR分析。

乙醇沉淀方法操作：

（1）估计DNA溶液体积。

（2）调节一价阳离子浓度。如果DNA溶液盐浓度高，则TE缓冲液（pH8.0）稀释；如果盐浓度低，则加入下表4－7所列的盐溶液。

表4－7 DNA沉淀盐浓度

盐溶液	用量	最后浓度
2.5mol/L 乙酸钠（pH5.2）	1/10体积	0.25mol/L
3.0mol/L 氯化钠（pH5.2）	1/10体积	0.3mol/L
2.0mol/L 乙酸钠（pH5.2）	1/10体积	2.0mol/L

（3）溶液充分混合后，加入2倍体积冰冷乙醇，充分混合。

（4）如果DNA量少或DNA长度较小（小于1kb），置－20℃冷却30～60min，沉淀DNA。

（5）置0℃12 000r/min离心10min。如果DNA浓度低或长度较短，则需加大离心力和延长离心时间。

（6）弃上清夜，离心管倒置于滤纸上使液体流干。用加样器吸掉挂在管壁上的液滴。置真空干燥或冻干机中1～2min或37℃放置30min，以除去残留上清液。

（7）沉淀的DNA用适量缓冲液溶解，如在37℃保温5min，DNA更容易溶解。

用乙醇不能完全沉淀很短的DNA分子（小于200bp）。向DNA溶液加$MgCl_2$至浓度为0.01mol/L，可明显提高小分子DNA回收率。

此外，还有三氯乙酸可以沉淀寡核苷酸，除去标记反应中的放射性核苷酸或核苷。当DNA溶液的浓度大于50mg/ml时，用聚乙二醇（PEG）沉淀法几乎能沉淀全部DNA。此法操作简便，不损坏DNA，费用低廉，很适合大量制备。聚乙二醇沉淀法能分离质粒DNA和染色体DNA。

4.3.2 仲丁醇浓缩法

用正丁醇或2－丁醇等有机溶剂抽提水溶液，有些水分子进入有机相，而DNA和其他溶质不进入有机相。经有机溶剂几次抽提后，DNA溶液体积大大减少，因此可用此法浓缩DNA。步骤如下：

（1）向样品中加入等体积的仲丁醇，充分混合。

（2）室温下1 200r/min离心5min，或1 600r/min离心1min。

（3）去上层仲丁醇相，重复步骤（1）～（3），至溶液体积达到所需体积为止。

（4）最后用水饱和的乙醚抽提DNA溶液2次，以除去残留的仲丁醇。通过蒸发去残留乙醚。

2-丁醇用量过多时，水全部进入有机相，将导致 DNA 沉淀。丁醇会破坏聚苯乙烯，抽提时必须在聚丙烯管中进行。

4.3.3 Microcon 100 法

其原理是利用超滤膜的不同孔径的微孔允许比孔径小的分子通过膜，流到超滤膜下面，而大于孔径的分子仍保留在超滤膜上面，使小分子与大分子分离开来。Microcon 100 截留分子量为 100 000，Centricon 30 截留分子量为 300 000。引物、水等小分子通过滤膜，目的片段（大于 100 000 或 300 000 DLT）保留在膜上，反向离心可以收集目的片段。

方法：

（1）将 Microcon 100 套在 1.5ml 离心管中（产品随带），将 PCR 扩增产物全部加在超滤管中；

（2）2 000r/min 离心 15min；

（3）去掉 1.5ml 离心管，将超滤管反套在另一干净的 1.5ml 离心管（又叫收集管）；

（4）4 000r/min 离心 2min，收集液体，可直接用于测序反应。

4.4 DNA 的贮存

在适宜条件（如 TE，pH8.0）下，纯净的 DNA 能长期贮存而无明显降解。DNA 一般溶于 TE 溶液，4℃ 保存。在 pH8.0 下，DNA 的脱氨基速度比在中性时慢。TE 溶液中的 EDTA 螯合二价阳离子，抑制 DNA 酶的活性，防止 DNA 降解。因为二价阳离子是大多数核酸酶的激活剂，而核酸酶又恰是降解 DNA 的主要因素。EDTA 还能抑制微生物生长，从而抑制核酸酶的形成。通常，在贮存 DNA 的溶液中加一滴氯仿，防止霉菌生长。

将 DNA 在上述缓冲液中冰冻贮存时，产生缺口的速度更慢；但解冻时又可产生新缺口，因此不宜反复冻融，不宜使用自动除霜的低温冰箱贮存 DNA。

Chelex-100 提取的 DNA 溶液应于 4℃ 保存，对于长期保存的，最好将上清液转移到一干净离心管中，除去 Chelex-100 颗粒，于 -20℃ 下保存。

4.5 DNA 纯化

在 DNA 提取过程中，一些 DNA 内切酶和 Taq 聚合酶抑制剂随着一起被提取出来，这些抑制剂的存在，会干扰随后的进一步分析，因此需要将它们从 DNA 溶液中除去。下面介绍部分纯化 DNA 方法。

4.5.1 低熔点琼脂糖凝胶电泳回收 DNA 法

低熔点琼脂糖由于多糖链上引入了羟乙基，使得琼脂糖在 30℃左右成胶，在 65℃左右熔化。利用这些性质，进行 DNA 纯化简化了从凝胶回收 DNA 的程序。

（1）灌制一含有适当浓度的低熔点琼脂糖凝胶。待凝胶冷却至室温后，将凝胶连同支承用的玻璃板一起转移至冷室的水平面上，以保证凝胶充分凝固。

（2）将 DNA 样品与加样缓冲液混合，加到凝胶加样孔内。于 4℃进行电泳以保证凝胶不会熔化。

注：一个给定大小的 DNA 分子在低熔点琼脂糖凝胶中迁移比在普通琼脂糖凝胶中得稍快。

(3) 用手提式长波紫外灯确定所需条带的位置，以减少因辐射引起的DNA损失。

紫外线有害健康，尤其是对眼睛。实验时，应戴护目镜。

(4) 用刀片切下含目的条带的琼脂糖，并切碎后转移至一个干净的一次性的塑料管里。

(5) 加约5倍体积的20mmol/L Tris－HCl (pH 8.0)、1mmol/L EDTA (pH 8.0) 至琼脂糖块中。盖好管盖，于65℃温育5min，熔化凝胶。

(6) 加等体积饱和酚，将混合液来回倒置混合20s。在20℃以4 000r/min离心10min，回收水相。界面的白色物质即是粉状的琼脂糖。再用酚－氯仿、氯仿各抽提1次。

(7) 将水相转移至离心管中，加入乙酸胺 (0.2mol/L) 和2倍体积的4℃无水乙醇。混合液在室温下放置10min。然后离心沉淀DNA与溶解DNA。

4.5.2. Centricon 30或Microcon 100法

纯化的原理与方法同4.3.3。

硅胶吸附法纯化成本低，回收率及纯度没有Microcon 100超滤法高，但后者成本高。

4.5.3 纯化试剂盒

市场上有各种商品纯化试剂盒，其大多利用吸附和超滤作用进行DNA的提取和纯化。正式用于检案前应按试剂盒说明书进行质量性能检测。

4.5.4 Sephadex G50过柱法

Sephadex G50过柱法是利用Sephadex的分子筛作用。

(1) 取一个0.5ml离心管，在管底和管盖上各扎一个小孔，孔径以针头斜面伸出一半为准。

(2) 吸取膨胀好的Sephadex G50液350μl，加到0.5ml离心管中，盖上管盖。

(3) 取一只1.5ml离心管，剪去管盖，把装有树脂的0.5ml管套在1.5ml离心管内放入离心机内，3 500r/min离心30s，吸出1.5ml管中的液体；再重复一次。

(4) 向小管内加入50μl TE缓冲液 (pH8.0)，同样条件离心，待加入TE缓冲液的体积与从大管中吸出的体积相同时，表明柱已平衡。

(5) 把待纯化的DNA溶液50μl加入小管，按平衡时的条件离心。

(6) 向回收的50μl层析液中加入NP－40、Tween－20、1mol/L $MgCl_2$、蛋白酶K使其终浓度分别为0.5%、0.5%、5mmol/L和100μg/ml，混匀后于37℃消化1h，沸水浴10min，取出备用。

应注意的事项是，本方法操作时必须使Sephadex G50柱平衡，否则会造成DNA的损失；本方法对DNA溶液有稀释作用，对于浓度很低的样品建议应在浓缩后应用。

第五章 DNA 定量、DNA 片段电泳分离与检测

5.1 DNA 的定性、定量分析

在任何 DNA 分析开始前，重要的是必需弄清有多少 DNA 存在，而且其中有多少是人 DNA，是否存在 DNA 降解，降解的程度又如何。对于 RFLP 来说，最终分析结果质量好坏主要取决于样品是否含有足够量的高分子量 DNA，而且要求了解样品的 DNA 含量，以便计算每个反应的酶量。对于 PCR 为基础的检验，模板 DNA（PCR 起始片段）的平均长度不太重要（尽管他们仍然必须长到有引物结合区域），但是 DNA 的量很关键。有时反应中加入太多的样品反而效果不理想，得不到 PCR 扩增。因此，不管应用哪种方法，在进行 DNA 分型前，都需要进行 DNA 的定性定量分析。

5.1.1 DNA 定性定量方法

1. 分光光度法：

组成 DNA 分子的碱基，均具有一定的吸收紫外线的特性，其最大吸收值在波长 250～270nm 之间（表 5-1）。这些碱基与戊糖、磷酸形成核苷酸后，其吸收紫外线的特性没有改变，核酸的最大吸收波长为 260nm。这个物理特性为测定 DNA 浓度提供了基础。

表 5-1 碱基的最大紫外吸收

碱基	腺嘌呤	胞嘧啶	鸟嘌呤	胸腺嘧啶
最大吸收值（nm）	260.5	267	267	264.5

在波长 260nm 紫外线下，1 OD 值的光密度相当于 50μg/ml 双链 DNA，40μg/ml 单链 DNA 或 RNA，20μg/ml 单链寡核苷酸。用 UV-260 型紫外分光光度计测定 260nm 处的吸收值，以此计算 DNA 溶液浓度。这种分光光度方法操作简便、快速，不但能确定 DNA 的浓度，还可以通过计算 260nm 和 280nm 的紫外吸收的比值（即 OD260/OD280 比值）粗略确定 DNA 的纯度，纯品 DNA 的 OD260/OD280 比值为 1.8～2.0，如果样品中污染有蛋白质或酚，其 OD260/OD280 比值将明显低于此值，此时无法对样品中 DNA 进行精确定量。该法不能确定样品降解、是否为人 DNA，不能确定样品含有多少高分子量 DNA。

操作步骤：

（1）取两只干净的比色杯，各加入 2ml 0.1mol/L NaOH，校正零点。

（2）以其中一只比色杯为空白对照，另一比色杯中加样品 DNA 溶液 4μl，加 0.1mol/L NaOH 至 2ml，混匀。

（3）测定波长为 260nm 时的 OD 值，再将波长调至 280nm，测其 OD 值。

（4）根据 $OD_{260}=1$ 时，双链 DNA 浓度为 50μg/ml，单链 DNA 或 RNA 浓度为 40μg/ml，计算样品 DNA 浓度。

计算公式：样品DNA浓度（μg/ml）＝OD_{260}×40μg/ml×稀释倍数

⑤计算OD_{260}与OD_{280}比值，若比值大于1.8时说明DNA样品比较纯。若比值小于1.6，则说明存在蛋白质或其他吸收紫外波长的杂质。

2. 琼脂糖凝胶电泳结合EB荧光测定法

有时，DNA的含量低（＜250ng/ml），无法用分光光度法进行测定，或有时DNA溶液中可能含有其他吸收紫外线的物质而影响DNA的精确定量。估算这些样品中DNA含量的一种快速方法是测定嵌入DNA中的溴化乙锭（EB）分子受紫外光激发而发射的荧光强度。由于这种荧光的强度与DNA总质量数成正比，通过比较样品与系列标准品的荧光强度，可对样品中的DNA进行定量。该法可检测出低至1～5ng DNA，可了解样品降解程度，确定有多少高分子量DNA。但该法同样不能确定是否为人DNA。

（1）制备0.8%的微型琼脂糖平板凝胶，含EB 0.5μg/ml。

（2）取2 μl DNA样品与0.4μl电泳上样缓冲液（含溴酚蓝）混合，混匀后加样于含EB的琼脂糖凝胶加样孔内。

（3）取一系列浓度不同的2μl标准DNA样品（0.5～50μg/ml）与0.4μl上样缓冲液混合上样。

（4）在80V电压下电泳约30min，使DNA移入琼脂糖胶内。用短波紫外线（254nm）进行拍照，比较样品DNA与DNA标准品的荧光强度，并计算出待测样品DNA浓度。

该法灵敏、快速，但不十分准确。它的荧光强度取决于嵌入碱基中的溴乙锭的多少，而且与DNA的螺旋程度密切相关，标准品与样品难以完全一致，并且还受其他发光污染物的影响。

3. 人特异探针滤条杂交法

利用人α－卫星基因座D17Z1探针进行杂交检测。在17号染色体上，α－卫星由500～1 000个重复单位组成，每个重复单位为2.7kb，多数由16个α单体组成，少数由15个单体组成，每个α单体为171bp。各个重复单位约98%碱基序列一致，各个α单体约65%～94%的碱基序列一致。α单体为人类特异。

通过滤条转移将样品DNA固定在尼龙膜上，生物素标记的或酶直接标记的D17Z1序列特异的探针与固定在尼龙膜上的样品DNA杂交，然后进行酶化学显色或化学发光法检测。比较待测样品与标准浓度DNA之间的感光强度或颜色深浅，确定待测样品浓度。

（1）将尼龙膜在水中湿润后，置10×SSC中浸泡10min。同时，将抽滤加样器用0.1mol/L NaOH清洗，然后用消毒水洗净。

（2）将两张预先用10×SSC湿润的滤纸铺在点样器的支撑物上，然后将湿润的尼龙膜放在滤纸上，小心排除气泡。安装好点样器，并接通真空室与真空泵。

（3）在加样孔中装满10×SSC，真空抽滤至所有液体被抽干。关闭真空泵。在加样孔中重新加满10×SSC。将样品孔中10×SSC抽干，关闭真空泵。

（4）样品的处理：DNA样品直接置沸水中变性10min，然后速置于冰水浴中，或取1μl待测样品及已知不同浓度的标准DNA与等体积样品变性液（0.5mol/L NaCl，0.5mol/L NaOH）充分混合，混合液于室温放置5min，使之充分变性为单链。

（5）将待测DNA与标准浓度DNA样品加到加样孔中，真空抽滤。待全部样品均被抽

干，加 10×SSC，抽滤，重复两次。

(6) 继续保持真空 5min，使尼龙膜干燥。

(7) 取下尼龙膜，室温下使其充分自然干燥。然后将尼龙膜放在用 0.4N NaOH 湿润的滤纸上固定 2min，DNA 面朝上。固定后的尼龙膜在 5×SSC 中漂洗 1min。

如果条件不允许，可采用手工直接点样。将 DNA 变性后，用微量加样器直接点样于干燥的硝酸纤维素膜上。注意必须采用小量多次法加样，待第一次样品完全干燥后，再在原来的位置第二次点样，使其直径不超过 1cm 为宜。

(8) 与人特异探针 D17Z1 杂交。探针杂交有碱性磷酸酶标记法和辣根过氧化物酶标记法等。下面列出碱性磷酸酶标记法（按美国 ACES 公司试剂盒方法）。

①转移好的尼龙膜放入杂交管中，倒入预杂交液，（预杂交液量为：$1ml/cm^2$），50℃预杂交 20min。将杂交液与 1.5μl 人类 DNA 探针混合后，置 50℃预热，备用。

②弃去预杂交液，加入预热好的含有探针的杂交液，50℃杂交 20min。

③用 1:5 稀释的洗膜液在 50℃洗膜两次，每次 10min。

④室温下用 20ml 终洗膜液洗膜两次，每次 5min。用滤纸吸干尼龙膜上的液体，将反应液（Lumi－Phos Plus）用加样器均匀加在膜上，室温反应 5min，然后用滤纸吸去多余的反应液（Lumi－Phos Plus），用保鲜膜包好检测后的膜，放入暗盒，压上 X 光片，在室温下使 X－光片感光 2 小时。冲洗 X 光片，比较待测样品与不同浓度已知标准的斑点感光强度，得出 DNA 浓度，计算 DNA 的含量。

此法灵敏度高，可检测微量 DNA，并且可以确定总 DNA 中含有多少人 DNA，其缺点是不能了解 DNA 的降解情况。

4. 人特异的 AluQuant™定量系统

(1) 原理　AluQuant™定量系统是利用与人基因组的高度重复系列特异的 DNA 探针，利用焦磷酸化和荧光素酶产生荧光，荧光仪测定发出的与 DNA 浓度相关的荧光强度，得到荧光强度与 DNA 含量关系的标准曲线，根据标准曲线获得待测样品的浓度。图 5－1 是其检测原理图，图 5－2 为标准曲线。

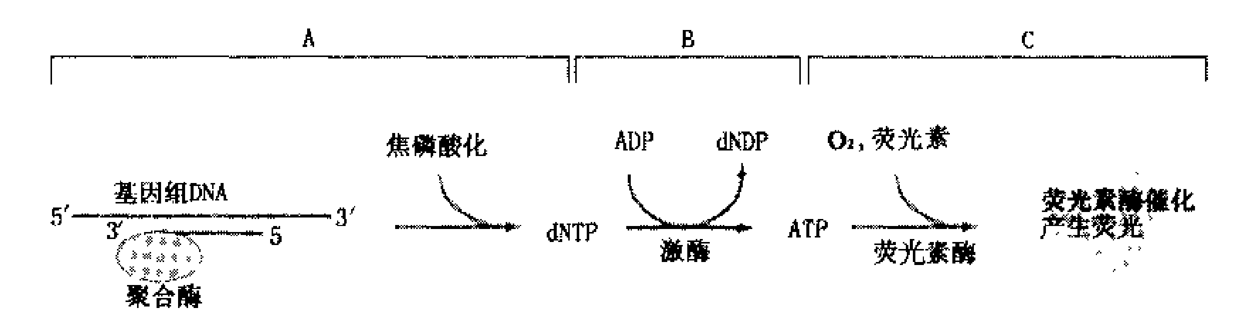


图 5－1　AluQuantTM 定量系统原理图

A：人 DNA 与探针序列退火形成杂交体，READase 识别杂交体，作为焦磷酸化的底物，进行焦磷酸化反应，双链 DNA 的 3′端加入一个焦磷酸，结果从 DNA 链释放出一个 3′端碱基，形成 dNTP。B：释放出的 dNTP 通过激酶将其磷酸转移到 ADP 中形成 ATP；C：形成的 ATP 在荧光素酶的作用下，与荧光素反应发出荧光，用荧光仪记录荧光信号。

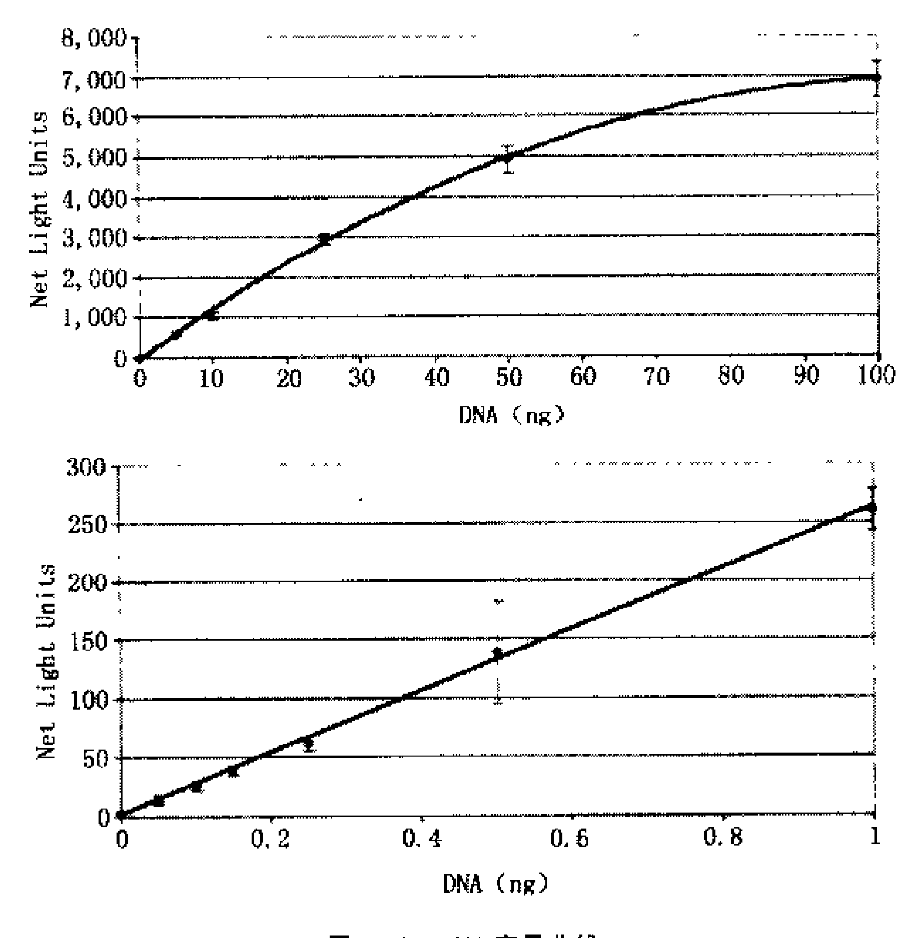


图5-2 DNA定量曲线

上图：DNA量为1～100ng的标准曲线；下图DNA量为0～1ng的标准曲线

该方法不涉及凝胶电泳与DNA的转移和固定，方法非常简单。可以准确测定0.1～50ng DNA，人DNA特异性比较高，果蝇、兔、鼠、鸡、牛、酵母及大肠杆菌的DNA基本上没有荧光信号，或非常低，与人DNA同样量的鼠、酵母及大肠杆菌DNA的荧光信号几乎接近零，而人DNA的荧光强度接近100%。另外，降解DNA并不影响DNA的定量分析。

(2) 方法

①试剂

标准DNA：用0.1×TE缓冲液稀释人DNA，制成浓度为4ng/μl，2ng/μl，0.5ng/μl，0.1ng/μl，0.02ng/μl和0ng/μl的一系列DNA溶液。

荧光素酶和荧光素试剂：将12ml的ENLITN荧光素酶/荧光素缓冲液加到ENLITN荧光素酶/荧光素瓶中，上下颠倒混合溶液，在室温放置60min以上，使其平衡。

Master mixes液的准备：每个分析准备二份Master mixes，一份含有AluQuant™探针混合物，另一份为没有核酸酶的纯水。Mixes含有5μl AluQuant™探针中和液，10μl酶溶液（含有READase聚合酶和READase激酶），5μl AluQuant™探针混合液或水。

②分析过程：各取二份5μl DNA样品，分别加5μl变性溶液，室温变性10min。在一份变性DNA溶液中，加入20μl含有探针的master mixes，另一份DNA溶液中加入没有探针的

master mixes。在 55℃下反应 60min。反应后，取 25μl 反应物转移到含有 50μl 的荧光素酶/荧光素的荧光管中。放置 3s。用 TD20/20 荧光仪检测荧光产量，测量的单位为相对荧光强度（relative light units，RLU）。

③计算 DNA 量：减去空白管的 RLU，用 macro 计算样品或 DNA 的净 RLU，获得净 RLU－DNA 浓度的标准曲线，然后将样品的 RLU 与标准曲线比较，计算样品 DNA 含量。

5.1.2 DNA 定性、定量方法的选择

选择什么样的 DNA 定性定量分析方法，既取决于原始样品的大小与质量，也取决于 DNA 的提取方法。比如，分析人员怀疑 DNA 量可能很少和/或已降解，可以用一小部分 DNA 进行滤条或斑点杂交；又如，用 Chelex 方法提取的 DNA，DNA 以单链形式存在，那么只能用斑点杂交法来检测；如果分析人员确信得到了大量 DNA 的话，则可以用凝胶电泳 EB 染色法来作定量分析，可以了解有关样品的降解程度。如果斑点杂交和凝胶电泳 EB 染色法一起进行，还可以了解有百分之几的是人 DNA，而不是细菌 DNA。有时，在凝胶电泳 EB 染色法后进行 Southern 杂交，用斑点杂交的人特异探针杂交检测 DNA 转移膜，也可帮助了解样品中人 DNA 的含量。因此，究竟采用什么样的定性定量方法，除案件情况和物证条件外，而且也得看分析人员认为哪些资料有用，来决定究竟选用哪种方法。个别的原始检材量太少了，可以先进行 PCR 分型分析，做完分析后反过来评估 DNA 的量，或用毛细管电泳进行定量定性测序。毛细管电泳极准确灵敏，可以测定极微量样品中 DNA 片段大小。但像凝胶电泳 EB 染色法一样，也不能区分是人 DNA 还是其他来源的 DNA。

5.2 DNA 片段平板凝胶电泳分离与检测

不同大小或不同碱基序列的 DNA 片段的分离，主要是利用电泳技术，其次，利用高压液相色谱或时间飞行激光离子解析质谱分析（MALID－TOF）。电泳是带电荷的物质在电场中的趋向运动。DNA 分子是一种强极性分子，等电点（pI）为 2.0～2.5，含有大量磷酸，在中性和碱性溶液中带负电荷，所以能在电场中向阳极移动。带电粒子在一定电场强度下，单位时间内在介质中的迁移距离叫迁移率，迁移率的大小与样品分子所带的电荷、电场中的电压及电流成正比，与样品的分子大小、介质粘度及电阻成反比。目前应用于 DNA 分析的有平板凝胶电泳和毛细管电泳。平板凝胶电泳根据凝胶基质不同，有琼脂糖和聚丙烯酰胺凝胶电泳。琼脂糖凝胶的孔径大，可以分离长度为 100bp～60kb 的 DNA 分子，聚丙烯酰胺凝胶的孔径小，分离较小片段（5～1 000bp）效果好。琼脂糖凝胶的分辨率比聚丙烯酰胺凝胶的低，但分离范围宽。

5.2.1 DNA 电泳分离机制

每个核苷酸含有一个磷酸基团，DNA 分子的荷质比是恒定的，因此一个含有 10 个核苷酸的 DNA 片段与另一个含有 100 个核苷酸的 DNA 分子在电场中受到的拉力是相同的。因此必须用筛分（分子筛）的机制分辨大小不同的 DNA 分子。凝胶或多聚溶液是一种含有一定大小孔径的基质，它像一个分子筛，较大的 DNA 分子在通过凝胶这些孔径时被滞留，而较小分子顺利通过凝胶孔，向前移动的速度比大分子 DNA 快，不同大小的分子彼此分离。DNA 分子通过凝胶机制有两种主要的模式：OGSTON 筛分（*ogston sieving*）与爬行（reptation）模式。

OGSTON 筛分模式认为 DNA 分子是一个圆球或像一个缠绕的线团，翻滚着通过凝胶孔，分子通过凝胶的速度与找到能让它们通过的足够大的孔的能力成正比。较小分子能通过大的孔，移动快，而当 DNA 分子大于凝胶筛分的孔径时，不能通过孔径，OGSTON 模式预测分子的迁移速度将变成零。

爬行模式认为 DNA 分子像蛇一样爬行通过凝胶，DNA 分子被拉长成一根线，进入胶的基质中，不同大小的 DNA 分子蜿蜒移动通过凝胶孔，按其分子量大小被分开。

电泳是一个相对的测量技术，如果不与已知 DNA 片段大小的标准物比较，凝胶中的 DNA 带位置是没有意义的。因此在电泳时要求在凝胶中同时加样分子量标准物。

5.2.2 影响 DNA 迁移率的因素

1. 凝胶浓度

一给定大小的线状 DNA 片段，其迁移率在不同的凝胶中各不相同。DNA 电泳迁移率（μ）的对数与凝胶浓度（τ）成线性关系，即

$$\lg\mu = \lg\mu_0 - k_r\tau$$

其中，μ_0 为 DNA 的自由电泳迁移率，k_r 为回归系数，是与凝胶的性质、迁移分子的形状和大小有关的常数。

选用不同的凝胶浓度可以分辨大小不同的 DNA 片段，表 5－2 列出不同凝胶浓度与其分离 DNA 分子的范围。

表 5－2　凝胶浓度和 DNA 分子的有效分离范围

凝胶	胶浓度（%）	线性 DNA 分子大小
聚丙烯酰胺	3.5	1 000～2 000（bp）
	5.0	80～1 000
	8.0	60～400
	12.0	40～200
	20.0	6～100
琼脂糖	0.3	5～60（kb）
	0.6	1～20
	0.7	0.8～10
	0.9	0.5～7
	1.2	0.4～6
	1.5	0.2～4
	2.0	1～3

2.DNA 分子的大小

线状双链 DNA 分子在一定浓度凝胶上电泳，电泳迁移率与碱基对数目的常用对数成

反比（图 5-3）。分子越大，在凝胶中移动越慢。

3. DNA 分子构象

分子量相同、构象不同的 DNA 如超螺旋环状（I 型）、带切口环状（II 型）和线状（III 型）DNA 通过凝胶时速度不同。它们的相对迁移率主要取决于凝胶的浓度，但也受电流强度、缓冲液的离子强度及 I 型 DNA 超螺旋程度的影响。在相同分子量情况下，高度螺旋 DNA 迁移率比线形 DNA 大。

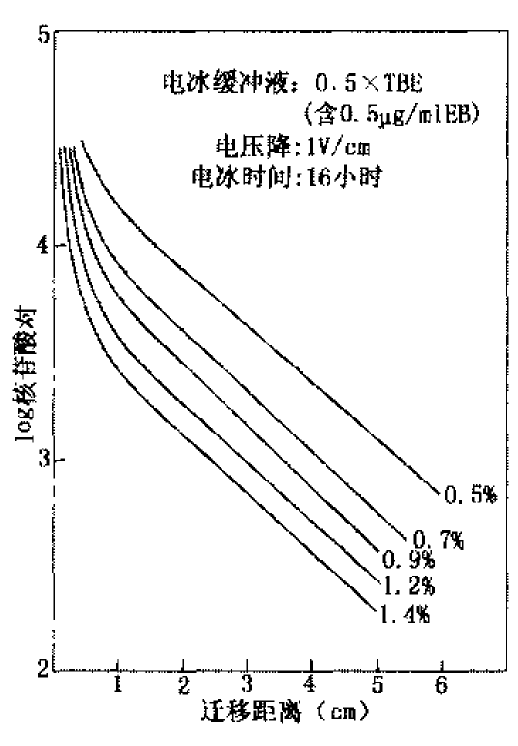


图 5-3 凝胶 DNA 迁移率与分子量关系

4. 电压

在低电压下，线形 DNA 片段的迁移率与所加的电压成正比。但随着电压升高，高分子量 DNA 片段的迁移率将以不同的幅度增大。因此，随着电压的增加，琼脂糖凝胶的有效分离范围缩小。要使大于 2kb 的 DNA 片段达到最佳分辨率，凝胶两端所加电压不要超过 5V/cm。

5. 电场方向

如果电场方向保持不变，则大于 50~100kb 的 DNA 分子在琼脂糖凝胶上的迁移速率相同。但是，如果电场方向呈周期性改变，则 DNA 分子被迫改变路径。由于 DNA 分子越大，为适应新的电场方向而重新排列所需的时间就越长，因此可以通过脉冲电场凝胶电泳来分离极大的 DNA 分子（达到 10 000kb）。

6. 碱基组成与温度

DNA 在琼脂糖凝胶中的电泳行为（与聚丙烯酰胺中截然不同）受 DNA 的碱基组成或凝胶电泳温度的影响不明显。在琼脂糖凝胶电泳中，不同大小的 DNA 片段的相对迁移率在 4℃与 30℃之间没有变化。因此，琼脂糖凝胶电泳一般可在室温下进行。但是，浓度低于 0.5%的琼脂糖凝胶和低熔点琼脂糖凝胶较为脆弱，最好在 4℃下电泳。在低温时它们强度较大。但 DNA 在聚丙烯酰胺凝胶中，DNA 分子迁移与碱基组成或温度有关。条件不同，可以分辨只相差一个碱基序列的二个 DNA 片段，如 SSCP 分析。

7. 嵌入染料的存在

荧光染料溴化乙锭用于检测琼脂糖和聚丙烯酰胺凝胶中的 DNA，它会使线状 DNA 的迁移率降低 15%。染料嵌入到堆积的碱基对之间，并拉长线状和带切口的环状 DNA，使其韧性更强。

8. 电泳缓冲液的组成

电泳缓冲液的组成及其离子强度影响 DNA 的电泳迁移率。电泳缓冲液的成分应避免与被分离的样品发生化学反应而改变样品的理化性质，或使其丧失生物活性。缓冲液离子强度一般选择在 0.02~0.2mol/L 范围。缓冲液中加 EDTA，以螯合金属离子，抑制 DNA 酶活性，保护 DNA。

在没有离子存在时（如用蒸馏水代替缓冲液配制凝胶），电导率最小，DNA 分子迁移

不动，即使 DNA 还能移动一点的话，也很像在高离子强度的缓冲液中（如误加了 10× 电泳缓冲液），电导很高并明显产热。最糟的是引起凝胶熔解，DNA 发生变性。

有几种不同的缓冲液可用于 DNA 的电泳（表 5-3）。这些缓冲液含有 EDTA（pH8.0）和 Tris-乙酸（TAE）、Tris-硼酸（TBE）或 Tris-磷酸（TPE）。其浓度约为 50mmol/L（pH7.5~7.8）。电泳缓冲液通常配制成浓缩液，贮存于室温，使用时稀释。

双链线状 DNA 片段在 TAE 中的迁移速率比在 TBE 或 TPE 中快 10%左右，但这些系统的分辨能力几乎相同，只是超螺旋 DNA 在 TAE 中的分辨率要比在 TBE 中更好。TAE 是使用最广泛的缓冲液，但它的缓冲容量相当低，长时间电泳会使其缓冲容量丧失殆尽（阳极呈碱性，阴极变成酸性）。在进行高压、长时间的电泳时，需更新缓冲液或在两槽之间加一蠕动泵进行缓冲液循环。TPE 和 TBE 比 TAE 成本稍高，但它们的缓冲容量明显较高。

表 5-3　常见电极缓冲液

名称	工作溶液	储藏母液
TAE	0.04mol/L Tris-醋酸， 0.002mol/L EDTA	50 倍：Tris 242g，冰醋酸 57.1ml， 0.5mol/L EDTA 100ml，pH8.0
TPE	0.08mol/L Tris-磷酸盐， 0.008mol/L EDTA	10 倍：Tris 108g，85%磷酸 15.5ml， 0.5mol/L EDTA 40ml，pH8.0
TBE	0.089mol/L Tris-硼酸， 0.002mol/L EDTA	10 倍：Tris 108g，硼酸 50 g， 0.5mol/L EDTA 40ml，pH8.0

5.2.3 DNA 电泳的指示剂与上样缓冲液*

5.2.3.1 指示剂

在 DNA 电泳中加一种带颜色的指示剂来指示样品的迁移距离，常用的指示剂有溴酚蓝（bromophenol blue，Bb）和二甲苯菁蓝（xylene cyanol，Xc）两种。

溴酚蓝在碱性溶液中呈蓝紫色，在中性溶液呈红棕色，在酸性溶液呈淡黄色。其分子量为 670，在不同浓度凝胶中，迁移速度不同，相当于不同分子量的 DNA。如在 0.65%、1%和 2%的琼脂糖凝胶电泳中，溴酚蓝的迁移率分别与 1kb、0.6kb 和 0.15kb 的双链线性 DNA 片段大致相同。在 5%的聚丙烯酰胺凝胶或含 7~8mol/L 尿素的 5%的聚丙烯酰胺凝胶电泳中，溴酚蓝分别相当于 65 和 35 个核苷酸的 DNA 片段。

二甲苯菁蓝为蓝色分子，在凝胶电泳中常为天蓝色，分子量为 554.6，携带的电荷量比溴酚蓝少很多，在凝胶中迁移比溴酚蓝慢。

所以根据指示剂的迁移情况，推断样品的电泳进程，决定电泳时间。表 5-4 是指示剂在不同凝胶浓度下相当的 DNA 大小。

表 5-4 指示剂在不同凝胶浓度下相当的 DNA 大小

	溴酚蓝（bp）	二甲苯青蓝（bp）
琼脂糖凝胶（%）		
1.0	1 000	2 000
1.4	200	1 600
聚丙烯酰胺（%）		
5	65	280
8	25	105
10	25～30	95
7mol/L 尿素聚丙烯酰胺（%）		
6	30	110
8	18	70
10	15	60
15	12	46
20	9	29

5.2.3.2 上样缓冲液

为了加样时使 DNA 样品能沉入凝胶加样孔底部，样品中加有可见颜色指示外，还要求加入适量的大分子如蔗糖、聚蔗糖和甘油等，以增加比重。目前常用的上样缓冲液如表 5-5所示。利用 DNA 测序仪分析时，由于使用高电压，要求样品中离子强度低，常用甲酰胺作为变性剂，使 DNA 分子保持单链，而且增加样品密度，利于加样，除甲酰胺外，对于平板凝胶电泳，为了加样方便，还加入一种蓝色的葡聚蓝，它是一种大分子，在电场中迁移慢，一直滞留在加样孔处，不干扰荧光检测。像 ABD 310 测序仪自动进样，就不需要加颜色指示剂。

表 5-5 DNA 上样缓冲液

缓冲液类型	6×缓冲液	应用	保存温度
Ⅰ	溴酚蓝 0.25%二甲苯青蓝 40%（W/V）蔗糖 或 4%（W/V）聚蔗糖 或 30% 甘油	琼脂糖凝胶、 非变性聚丙烯酰胺凝胶	4℃
Ⅱ	0.25%溴酚蓝 0.25%二甲苯青蓝 40%（W/V）蔗糖 95%甲酰胺	变性聚丙烯酰胺凝胶	4℃
Ⅲ	葡聚蓝 95%甲酰胺	377 等测序仪分析	4℃

5.2.4 分子量标准

同一凝胶中，迁移距离相同，则分子量相同。比对待测 DNA 片段与分子量标准的电泳迁移距离，可以确定 DNA 片段大小。常用的分子量标准见表 5-6。

表 5-6 DNA片段分子量标准（bp）

λHindⅢ	λHindⅢ - EcoRⅠ	EcoRⅠ	pUC18 - Sau3AⅠ	φX174 - HaeⅢ
23 130	21 226	21 226	955	1 353
9 416	5 148	7 421	585	1 078
6 587	4 973	5 804	341	872
4 361	4 268	5 643	258	603
2 322	3 530	4 878	141	310
2 027	2 027	3 530	105	281
564	1 904		78	271
125	1 584		75	234
	1 375		46	194
	947		36	118
	831			72
	564			
	125			

5.2.5 琼脂糖凝胶电泳

琼脂糖是从海藻中提取出来的一种线状高聚物，其基本结构为：

HO　CH₂OH　HO　HO
D-半乳糖　3,6-脱水-L-半乳糖糖

琼脂糖有多种等级，即使在同一等级中，不同批号间质量也有明显的差异；其他多糖、盐及蛋白质的污染程度不同，这些差异影响DNA的迁移。根据经验，针对一般目的最好使用低电渗琼脂糖，它容易融化产生透明的溶液，所形成的凝胶即使在低浓度时仍能保持一定的机械强度。

1. 电泳槽

琼脂糖凝胶大多采用水平潜水电泳(图5-4)。这种系统有4个优点：(1) 整块凝胶下面有支撑，可使用低浓度琼脂糖；(2) 可以制备不同大小的凝胶；(3) 易于灌胶制板、加样和操作；(4) 仪器耐用，造价低。

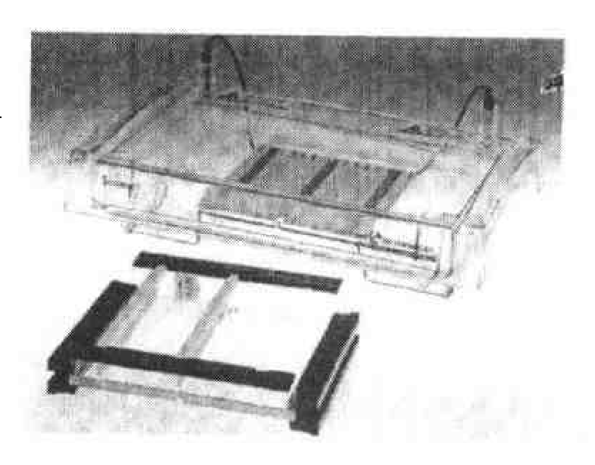


图 5-4 潜水式电泳

2. 凝胶制备与电泳

(1) 根据浓度称取所需粉状琼脂糖，加到500ml三角瓶中，加入200ml 1×TAE缓冲液，配制200ml琼脂糖凝胶。用记号笔在三角瓶的外表标记液面高度，用来指示加热蒸发后需加多少蒸馏水。

(2) 在沸水浴或微波炉中加热直至琼脂糖完全溶解，加蒸馏水补充至原体积（200ml）。

(3) 将溶液冷却至50℃，加入EB（10mg/ml，4℃避光保存）使终浓度为0.5μg/ml。

(4) 将干净玻璃板放在制胶盒中，用胶带或有机玻璃挡板把制胶盒两端围住。然后放在水平台上，用水平珠调平。

(5) 用吸管取少量琼脂糖溶液封固制胶盒四周。

(6) 封固后室温放置5~10min，距制胶盒一端约1cm处插入加样梳子，使梳齿在靠近胶一端的位置形成加样孔，调节梳子高度，使齿底与玻璃板间距离为0.5~1.0mm，然后把剩余的琼脂糖溶液倒入制胶盒。

(7) 凝胶在室温下凝固30~45min，再置4℃下放置10min。小心拔出梳子，去除两端的胶带或挡板，将凝胶放入电泳槽。

(8) 在电泳槽两端加1×TAE，使缓冲液液面高出胶面约5mm。电极缓冲液应与凝胶缓冲液一致，如果电泳缓冲液与凝胶中的离子强度不一致将影响DNA片段迁移。或先加电极缓冲液至胶面，将DNA样品加到凝胶中，电泳5min让DNA进入凝胶中后，再加剩余的电极缓冲液，使溶液液面高于凝胶面约5mm。

(9) DNA液中加1/10体积的10×载样缓冲液，混合，快速离心15s，使溶液全部汇集管底部。

(10) 加样前要用吸头吹打凝胶孔中的溶液以赶走样品孔中的气泡，将DNA样品加入到凝胶加样孔中。DNA分子量标准加在样品旁边。加样时吸头不要碰坏凝胶孔壁，否则DNA的带型不整齐。每加样一个样品，更换1个吸头，避免样品交叉污染。

DNA上样量取决于DNA样品片段长度、数目及加样孔形状与容量。对于0.5cm宽的加样孔，电泳后每条DNA带最低检测量为2ng。如果每条DNA带500ng，会发生带型拖尾、轮廓不清的超载现象。酶切后基因组DNA，含有很多DNA片段，上样20~30μg会明显地影响分辨率。

对于标准的0.5cm×0.5cm×0.15cm的样品孔，最大上样体积为37.5μl。若样品体积大于凝胶孔体积，会导致样品溢出，流入邻近样品孔造成样品交叉污染，影响结果分析。如果上样量多，可将凝胶制作厚一点，增加上样孔的体积，或者将样品浓缩成较少溶液体积。

(11) 电源接通前核实正、负电极是否放置正确。

(12) 接通电流，观察正、负极铂金丝是否有气泡发生，如负极的气泡（H_2）比正极的气泡（O_2）多一倍，则表示电泳槽已经接通电源，几分钟后可见到指示剂溴酚蓝向正极移动。

(13) 用万用表测量并调整凝胶端电压降为5V/cm。因为电泳仪上数字显示的输出电压，不一定是电泳槽两极之间的实际电压，以万用表测量凝胶两端电压为准。

(14) 根据溴酚蓝迁移情况，切断电源、取出凝胶，进行下步操作。含EB的凝胶可以直接于紫外灯下观察结果或拍照记录。

加溴乙锭的凝胶在电泳过程中，溴乙锭向负极移动，与DNA的泳动方向相反，较长时间的电泳会造成靠正极方向的凝胶中EB含量很低，导致小分子量DNA片段检测困难，需在电泳后重新将凝胶在0.5μg/ml EB溶液中染色30~45min。

凝胶中加入EB进行电泳，便于随时在紫外灯下观察电泳状态。但EB会导致线性

DNA迁移率下降15%，并可能影响DNA带型的整齐，为了获得整齐、狭窄的DNA带型，最好在电泳后进行EB染色。

5.2.6 **聚丙烯酰胺凝胶电泳**

1. 概述

丙烯酰胺是一种单体，在催化剂N，N，N'，N'－四甲基乙二胺（TEMED）和过硫酸胺的存在下，丙烯酰胺聚合，形成长链，加入交联剂N，N'－亚甲双丙烯酰胺（Bis），聚丙酰胺链就会交叉连接形成凝胶。凝胶的网孔大小（孔径）取决于聚合链的长度与交联的程度（交联度）。聚合链的长度取决于聚合反应中的单体丙烯酰胺（Acr）的浓度，交联度取决于交联剂（Bis）的含量。调节单体与交联剂的浓度比例，可形成不同交联度的网状聚合物。凝胶浓度和交联度的计算公式如下：

$$T\% = \frac{(x+y) \times 100}{V} \qquad C\% = \frac{y \times 100}{x+y}$$

式中，x、y分别为单体丙烯酰胺和N，N'－亚甲双丙烯酰胺的量，V为体积数。

交联剂浓度越大，网孔孔径越小，适于DNA小片段分离。不同浓度非变性丙烯酰胺凝胶的分离能力见前表5－2。

聚丙烯酰胺凝胶与琼脂糖凝胶相比有4个主要优点：①分辨率高，500bp以下的仅仅相差1bp的DNA分子即可分开；②所能装载的DNA量远远大于琼脂糖凝胶：多达10μg的DNA可以加样于聚丙烯酰胺凝胶的一个标准样品槽（10mm×1mm）而不致显著影响分辨率；③从聚丙烯酰胺凝胶中回收的DNA纯度很高；④聚丙烯酰胺凝胶无色透明，比琼脂糖凝胶的紫外线吸收低，抗腐蚀性强，机械强度高，韧性好。常用于PCR扩增产物的分离检测。

常用的聚丙烯酰胺凝胶有两种：

（1）非变性聚丙烯酰胺凝胶　用于分离和纯化双链DNA片段。凝胶缓冲液为1×TBE，在低电压（1～8V/cm）下电泳，以防止电流通过时产生的热量引起小片段DNA的变化。大多数双链DNA在非变性聚丙烯酰胺凝胶中的迁移速率大略与其大小的lg对数值成反比，但迁移率也受其碱基组成和序列的影响，以致大小完全相同的DNA其迁移率相差可达10%。这是由于双链DNA在特定序列上形成扭结造成的。由于不可能知道一种未知DNA的迁移是否规则，所以非变性聚丙烯酰胺凝胶电泳不能用于双链DNA的大小确定。

（2）变性聚丙烯酰胺凝胶　用于分离、纯化单链DNA。凝胶中加入6～7mol/L尿素作为变性剂，阻止DNA中的碱基配对。变性DNA在变性聚丙烯酰胺凝胶中的迁移速率几乎与其碱基组成及序列完全无关，只与DNA分子的大小有关。STR分析和DNA测序反应产物的分析均采用变性聚丙烯酰胺凝胶。

聚丙烯酰胺凝胶的厚度一般在1mm以下，常为0.4～0.75mm，过厚的凝胶因产生热而导致DNA电泳的带型不整齐，因此凝胶的厚度越薄越利于凝胶的分辨率，但太薄的凝胶不太好操作，机械强度低。根据分离检测的DNA片段长度的需要，凝胶长度可以在10～100cm之间。对于只是为了确定是否有DNA，可以采用微型聚丙烯酰胺凝胶；为了分离DNA，尤其是要求区分差异较小的DNA片段，如STR－PCR产物的检测，应采用长距离凝胶电泳，如20cm～40cm长的凝胶。聚丙烯酰胺凝胶一般是进行垂直电泳，聚丙烯酰胺凝胶的制备和电泳都比琼脂糖凝胶更为费事。

2. 影响凝胶聚合速度的因素

凝胶聚合过快或过慢，会影响电泳分离效果。聚丙烯酰胺凝胶的聚合过程受多种因素影响。如氧气能淬灭自由基，延缓凝胶聚合，因此在丙烯酰胺单体液中加入引发剂之前常采用负压抽气处理。聚丙烯酰胺凝胶几乎总是铺于两块玻璃板之间，两块玻璃板由夹条（spacer）隔开并封以绝缘胶布，丙烯酰胺溶液不会与空气接触，所以氧对聚合的抑制仅限于凝胶顶部的一个窄层里。

凝胶聚合的最佳温度为 23 ~ 25℃，灌胶前要注意将丙烯酰胺单体贮存液、灌注凝胶的模具置室温下平衡至室温。降低温度明显减慢凝胶聚合，温度下降 1℃，聚合时间约延长 2 分钟。同样制胶条件下，夏天丙烯酰胺聚合比冬天快，在冬天为了加快胶的聚合，可以适当增加过硫酸铵和 TEMED 用量，或将胶放入 37℃的温箱内。在夏天可以减少过硫酸铵和 TEMED 用量，降低聚合速度。

引发剂和催化剂的用量决定凝胶聚合的速度，也影响凝胶的质量。增加过硫酸铵或四甲基乙二胺用量，将使丙烯酰胺聚合链长度缩短，凝胶变得脆弱而失去柔性。引发剂用量再增加至一定程度时，聚合链长度过短，凝胶聚合过程反受抑制。引发剂过少则聚胶缓慢，大气氧越来越多地进入单体溶液抑制聚合过程，结果胶孔增大，支撑力弱而易于变形。因此，通过改变引发剂，催化剂用量控制聚胶速度，保证完全聚胶。过硫酸铵浓度一般控制在 0.02% ~ 0.5%之间，使聚合在 30 ~ 40min 内完成。在加样前预电泳 15 ~ 30min 可使凝胶中残存的硫酸根离子泳出凝胶，减少引发剂过硫酸铵的影响。

pH 值对聚胶的影响与引发剂有关，过硫酸铵作引发剂，碱性条件下聚合快，酸性条件下聚合慢，例如 T7%凝胶在 pH8.8 时聚合时间约需 30min，在 pH3.3 时聚合要 90min。

灌制凝胶所用的试剂均需分析纯级，其中丙烯酰胺的纯度最重要。丙烯酰胺中出现的杂质会直接影响凝胶质量。丙烯酸是丙烯酰胺脱氨的产物，固体丙烯酰胺中可能含有丙烯酸，单体贮存液放置时间过长，部分丙烯酰胺将水解成丙烯酸。丙烯酸也能与丙烯酰胺，亚甲基双丙烯酰胺发生共聚合，结果使部分的凝胶区域的 pH 值发生改变。丙烯酰胺试剂中还有金属离子等其他杂质，都会引起凝胶聚合不良或聚合不均匀等。理想的丙烯酰胺试剂中丙烯酸含量应在 0.05%以下，配成 5%的溶液，pH 值不应低于 5。

TEMED 极易被氧化而失去催化能力，试剂中若污染有氧化剂杂质，溶液呈黄色。如果无法得到无色透明的试剂，可适当增加用量以保证凝胶聚合速度。过硫酸铵容易吸潮，溶解于水后很快降解并失去引发胶聚合的作用。所以过硫酸铵应避光，干燥，密封条件下保存，过硫酸铵溶液以当天配制为好，4℃条件下保存也不宜超过一周。

各类凝胶缓冲液成分对凝胶聚合无影响，如尿素不干扰聚丙烯酰胺凝胶的聚合。

3. 基本技术

垂直电泳槽装置见图 5 - 5。灌胶模具由两块玻璃组成，两侧边各有一厚 0.4 cm 或 0.8cm 的夹条，夹条的厚度即凝胶的厚度。凝胶液灌注一定高度后，插入加样梳，就可制成加样槽孔格，加样槽的大小由加样梳齿的宽度决定。

（1）试剂准备：丙烯酰胺贮备液浓度 T30%（Acr∶Bis = 29∶1），电泳缓冲液为 1 × TBE 或 0.5 × TBE，凝胶多采用过硫酸铵 - TEMED 系统。

（2）制胶：准备好合适的玻璃板、夹条、梳子。玻璃板要清洗干净以防灌胶时产生气

泡，从边缘拿取以免弄脏玻璃板的工作区。制胶模具两块玻璃间的两长边各夹一长夹条，两侧边用强力夹夹紧。根据模具的体积按照表 5－7 计算所需的配置胶液的用量。丙烯酰胺、亚甲基双丙烯酰胺、TBE 及超纯水等加入充分混匀后，加入 TEMED 和 10% 过硫酸胺溶液，混匀。用加样器小心吸取上述溶液，加入到两玻璃板间的空隙处，直至离顶部 0.3cm 左右。立即插入加样梳，室温下静置 30～60min。小心地取出加样梳，用水冲洗加样孔。

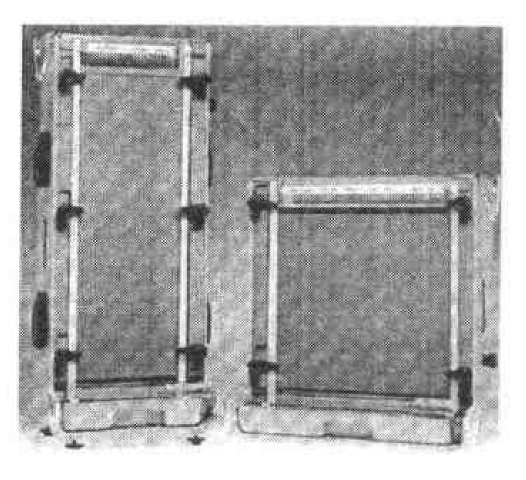
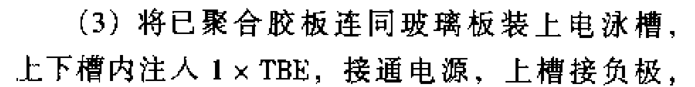


图 5－5 垂直凝胶电泳



（3）将已聚合胶板连同玻璃板装上电泳槽，上下槽内注入 1×TBE，接通电源，上槽接负极，下槽接正极，按 5～10V/cm 场强设置电压。根据载样缓冲液染料迁移部位确定样品 DNA 相对位置，终止电泳。

（4）取下模具及凝胶，用刀片撬下上层玻璃，进行 DNA 染色。

表 5－7 聚丙烯酰胺凝胶配制（每 100ml 胶液配方）

试剂	不同浓度（%）的凝胶各成分用量（ml）				
	3.5	5.0	8.0	12.0	20.0
30% 丙烯酰胺	11.6	16.6	26.6	40.0	66.6
H_2O	67.7	62.7	52.7	39.3	12.7
5×TBE	20.0	20.0	20.0	20.0	20.0
10% 过硫酸铵	0.7	0.7	0.7	0.7	0.7
TEMED	35μl	35μl	35μl	35μl	35μl

电极缓冲液可以反复使用几次，只是后来电极缓冲液的离子强度增加，电流要比新的缓冲液大。

垂直平板凝胶电泳制备涉及玻璃板的清洗、准备制胶模具、配制胶液、灌制凝胶以及固化等，过程冗长繁杂，劳动强度大、费时；丙烯酰胺具有极强的神经毒性，通过皮肤吸收；平板凝胶电泳上样时手动加样，必须十分小心，以免泳道间互相干扰，影响结果分析，而且工作强度大。

5.2.7 DNA 片段的检测（显现）

DNA 片段是无色无味的分子，肉眼无法观察到，需要通过一些手段显现后才能观察到。显现的方法主要有：（1）结合荧光溴化乙锭；（2）银染；（3）放射性同位素标记；(4）非同位素标记；（5）掺入荧光素标记。前二种显色方法，DNA 分子在电泳前不需要经过任何处理；后面的三种方法在电泳前需对 DNA 分子经过特殊标记处理。

1. 溴化乙锭

溴（化）乙锭（ethidium bromide，EB）是一种荧光物质，带有苯环结构，这种偏平分

子可以嵌入 DNA 双链的配对碱基之间，在紫外线激发下，发出红色荧光。激发荧光能量来源于两个方面，一是 DNA 吸收波长 260nm 的紫外线后将能量传递给 EB，二是结合在 DNA 分子中的 EB 本身，主要吸收波长为 300nm 和 360nm 的紫外线的能量，来源于这两方面的能量，最终激发 EB 发射出波长为 590nm 的可见光谱红橙区的红色荧光。EB - DNA 复合物中的 EB 发出的荧光，比游离的凝胶中的 EB 本身发射的荧光强度大 10 倍，因此不需要洗净背景就能较清楚地观察到 DNA 的电泳谱带。若 DNA 量很少、而 EB 的背景太深导致谱带看不清楚，可将凝胶浸泡于 1mmol/L $MgSO_4$ 中 1h 或 10mmol/L $MgCl_2$ 中 5min，使非结合的 EB 褪色，减少未结合的 EB 产生的背景荧光，这样可检测到 10ng 的 DNA 样品。

一般采用在凝胶或电泳缓冲液中加入终浓度为 0.5μg/ml EB 的染色方法，在电泳过程中可随时观察 DNA 的迁移情况。但是由于 EB 带正电荷，中和 DNA 分子的负电荷，同时它的嵌入增加了 DNA 分子的刚性，使迁移率减慢，因此不宜用于检测 DNA 分子大小的凝胶电泳。利用凝胶电泳比较或测定样品 DNA 含量，也不宜在凝胶中直接加入 EB，因为游离的 EB 分子向负极泳动会使样品中前后各条带染色不均匀，影响定量。因此应在电泳后将凝胶浸入 0.5μg/ml 的 EB 水溶液中 10min 进行染色。

EB 见光易分解，应放在棕色试剂瓶中于 4℃条件下保存。染色时也应避光。电泳后应立即染色观察或拍照记录。若不能马上拍照可用保鲜膜或塑料袋封好，在 4℃保存。单链 DNA 分子中常存在自身配对的双链区，也可嵌入 EB 分子，但嵌入的量少，荧光较弱，其最低检出量为 0.1μg。

不同波长的紫外线光源激发 EB 产色荧光的灵敏度不同，短波 254nm 的紫外线是由 DNA 吸收后传递给 EB，灵敏度较高，对 DNA 的损害作用最大，会造成 DNA 链断裂或形成嘧啶二聚体，照射时间过长还会引起褪色效应。300nm 和 360nm 的紫外线主要由 EB 分子直接吸收，对 DNA 的断链和二聚体形成影响较小，褪色效应轻。根据不同目的要求，可选用不同波长的光源（表 5 - 8）。

表 5 - 8 不同波长的光源与 EB 显色效果

波长(nm)	灵敏度	光损伤效应		褪色效应
		产生断裂缺口	形成嘧啶二聚体	
254	较高	多	多	严重
300	最高，是 254nm 的 1.5 倍	较少，是 254nm 的 1/15	是 254nm 的 1/84	较微
366	低，是 254nm 的 1/5	最少，是 254nm 的 1/75	无	几乎不褪色

也可以用 SYBR Green 代替 EB。SYBR Green 是一种新型荧光染料，激发波长为 490nm，发射波长为 520nm，灵敏度更高，染色时间可缩短到 5min。

2. 银染

Ag^+ 与 DNA 形成稳定复合物，然后甲醛使 Ag^+ 还原成银颗粒。硝酸银等试剂可使聚丙烯酰胺凝胶上的单链、双链 DNA 及 RNA 都染成黑（棕）褐色。

银染法的灵敏度比 EB 染色高 200 倍左右，在小于 0.5mm 厚的凝胶中，能检测出 0.5ng 的 RNA。其缺点是特异性不强，也能染色蛋白质，并和去污剂反应产生类似 DNA 被染成的颜色。染色的深浅与 DNA 碱性组成有关，与核酸的含量不成正比，定量不准确。

银与 DNA 稳定结合对 DNA 有破坏作用，不适于 DNA 回收的制备型电泳。

银染方法：用固定液（10%乙酸，5%乙醇）室温固定 10min 后，小心倒去固定液，去离子水洗凝胶两次，每次数秒；用染色液（12mmol/L $AgNO_3$）室温染色 30min，期间要不断摇动至凝胶沉入溶液中；小心倒去染色液，用去离子水洗凝胶两次，每次数秒；再用显色液（2.65%Na_2CO_3，0.04%甲醛）显色 5~10min，其间不断摇动，直至看到清晰的谱带；最后用 10%乙酸溶液停显，凝胶照相记录结果，干胶保存。

银染过程中，不能用手直接接触凝胶，应始终戴一次性手套操作，每更换一种溶液，需更换一次手套，否则易使凝胶产生高背景；在倒去溶液或用水洗胶时要小心，可用戴手套的手指压住胶一端，防止凝胶滑入水池；万一胶滑入水池，用多次超纯水洗凝胶，冲去沾在胶上的杂质，尽可能降低背景。

凝胶染色过程中所有试剂均应用超纯水（18MΩ）配制，溶液中不能含有 Cl^-，如果含有 Cl^-，凝胶变得不透明、谱带等变成乳白色。10%乙酸固定后不充分洗去乙酸，银染显色后带强度弱，或有时加入停显液后凝胶迅速由淡黄色转成桔红色，看不到谱带。因此一定要把乙酸溶液冲洗干净。

显色反应需在碱性环境中，1%NaOH 可以代替 Na_2CO_3，使显色反应加快，但有时也会增加背景颜色。如果显色反应在酸性环境中，则 Ag^+ 不能被还原，DNA 谱带显现不出来。

固定液和 $AgNO_3$ 溶液可以反复使用多次，只是染色时间要适当延长，灵敏度下降。

3. 放射性同位素与酶化学反应检测法

利用标记技术将放射性同位素如 P^{32} 或生物素直接标在底物或反应的引物上，使 PCR 扩增产物带上放射性同位素或生物素标记，随后根据所带的标记进行自显影或酶化学反应检测 DNA 片段；或者是转移在尼龙膜上 DNA 与带同位素或非同位素标记探针通过杂交结合，进行检测。后者详细内容可以参见分子杂交部分。

4. 荧光标记检测法

荧光染料由于含有共轭双键易吸收激发光，被激发到高能态（激发态），但激发态极不稳定，当跃迁回到基态时，多余的能量以发光的形式释放出来。用仪器如光谱仪、荧光扫描仪可以记录发出的光信号。将荧光染料标记在 DNA 片段，根据片段所带的特征荧光受激发后发出的荧光信号可以检测 DNA 片段。关于荧光标记的更多内容参见 STR 章节。

5.3 毛细管电泳

5.3.1 原理与基本特点

毛细管电泳（capillary electrophoresis，CE）是 80 年代后期迅速发展起来的一项新的电泳分离技术，是在毛细管内进行电泳。它与平板凝胶电泳相比，具有以下六个方面的特点：一是 DNA 样品加样、分离和检测全部自动化；二是高效、快速，毛细管散热快，可以应用高压电源分离，DNA 片段分离时间大大缩短，由平板电泳的数小时变为毛细管电泳的几分钟；三是微量，样品用量极少，利于样品重复检测；四是定量准确；五是每次每道毛细管只分离一个样品，分离完一个样品后重新自动灌注新的胶液分离下一个样品，不需要追踪校正样品的电泳泳道，也没有因样品间由于加样不当引起的交叉污染；六是样品分离与检测一体完成，不需要扫描凝胶图。

毛细管电泳的缺点：一是通量问题，单道毛细管一次只能分析一个样品，不适合大规模的分析，多道毛细管电泳可以克服这点，但多道毛细管电泳带来的问题是造价高；二是就目前的仪器而言，一台机器上不太适合同时进行基因分型和测序，因为 STR 分析与测序的毛细管长度和胶液型号不同，它们不兼容，不能通用。如果要交替进行 STR 分型和测序，不断更换毛细管和胶液，这种操作对毛细管的寿命不利，也浪费胶液。

5.3.2 毛细管电泳的主要设备及功能

毛细管电泳设备由石英毛细管、样品进样系统、高压电源和检测系统四部分组成（图 5-6）。高压电源包括二个缓冲液槽和连接高压的二个电极，检测系统包括激光激发器和荧光检测器及负责进样和数据收集的计算机系统。

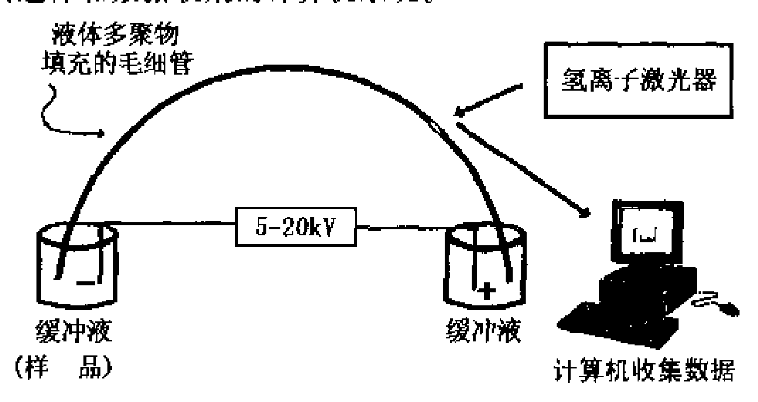


图 5-6　毛细管电泳示意图

毛细管是其十分重要的部件，一般内经为 50～100μm，长 20～100cm。ABI310 的用于 STR 基因分型的毛细管内径为 50μm，长为 47cm，用于测序的毛细管为 61cm。材料为熔融石英毛细管，石英材料物理化学性质稳定，有良好的传热、散热功能，对紫外线无吸收作用。管外涂有一层聚亚胺保护层，使毛细管具有一定的柔性而不易折断。涂层不透光妨碍毛细管内的检测，因此在接触激光激发部位的毛细管去掉涂层、暴露出毛细管。毛细管内壁一般没有保护层，直接与缓冲液、溶质或凝胶介质接触。与常规的凝胶电泳比较，毛细管电泳的最大优势是散热快。由于毛细管内径小，表面积与体积比大，易散热，保持分离的高效；内径细，电泳阻力大，加高场强仍然可维持较小电流，减少自热。毛细管高场强使毛细管电泳有高分辨率及缩短检测时间的明显优点。但内径越细，会增加毛细管的侧面积与体积之比，不利于吸附的抑制，而且又会造成进样、检测和清洗的困难。

影响毛细管长度的两个因素是峰的分辨率与分离时间。一般，毛细管越长，分辨率越好，分离时间越长。

对于毛细管电泳，有两个测量长度是十分重要的，一是检测长度（Ldet），指的是从样品加样口到激光检测口的距离，另一个是毛细管总长（Ltot），即样品从进样到出口的距离。改变 Ltot，影响 CE 系统的电场强度，而减短 Ldet，提高分离速度。理论上，最佳点是 Ldet 与 Ltot 越接近越好，但是仪器空间限制 Ldet 要比 Ltot 短 7cm 或 7cm 多。主要原因是电泳电极不可能与检测器挨得太近，ABD310 上检测窗口固定在离缓冲液出口 11cm 处，因此 Ldet 为 30 cm，PE 公司配有二种毛细管 47cm 和 61cm，短的那种用于 STR 基因分型等

基因扫描，对分离速度的要求比分辨率更高，能以极快的速度分离片段。较长的毛细管对DNA 测序更合适，因为对于分离只相差 1bp 的两较大片段，高的分辨率是首位，而对于分离时间的要求只是次要的。

ABI310 中的毛细管的内层有或没有共价结合的内层保护层。许多用于毛细管电泳的毛细管由化学衍生物包埋，阻止电渗流（electro - osmotic flow，EOF）。当 pH 大于 6 时，石英带负电荷，熔融石英的毛细管内层带负电，来自毛细管内的溶液的正电荷产生 EOF，沿负电荷的毛细管壁形成双层，在电场作用下，双层中的阳离子向阴极（进入口）移动，带动可溶分子向同一方向移动。

EOF 在另一相反的方向，与 DNA 分子的移动一致。EOF 在很大程度上取决于环境参数，比如 pH，温度、电压和缓冲液的粘度。因此，为了保证 DNA 片段分离的重复性，必须包埋毛细管的内壁以除去 EOF。有两种主要包埋方法：带电的硅烷基团的化学衍生或者粘性多聚体溶液的线性包埋。ABI310 毛细管电泳系统的毛细管一般内壁没有用硅烷基团包埋，而是用能进行 DNA 分离的 POP - 4 多聚体溶液线性包埋毛细管内层，防止 EOF。

样品进样系统有两种，电迁移（或称电动力）进样和水压式注入。电（迁移）进样方式：毛细管先不和缓冲液接触，而直接置于样品溶液中，然后在很短时间内施加进样电压，利用电场力的作用，样品中各组分依其离子电荷和迁移率不同有选择性地注入毛细管内。流体动力学（水压）式是将毛细管的注入端浸入样品液中，施加压力或真空力，样品溶液无选择性地进入毛细管。目前的仪器上都配有自动进样装置，实验人员通过进样时间与样品的浓度决定进样量。310 基因分析仪是利用电动进样模式进样。

毛细管电泳的高压电源可提供 15 kv 和 30kv，200 ~ 250mA 的直流电，并能根据需要改变正、负极的方向。电泳时一般采用恒压输出条件，但如果电泳时恒温控制不理想，可改用恒流电泳。因为温度控制是毛细管电泳的基本条件，如散热不良，毛细管内温度上升 1℃，会引起电泳缓冲液粘度出现 2% ~ 3% 的差异，进而影响样品的迁移率。

毛细管电泳的检测采用柱上检测，检测系统利用紫外、荧光和电化学等检测器，紫外检测器和荧光检测器是目前使用最广的两种检测手段。检测器一般位于毛细管的远端 4/5 处，紫外检测器的灵敏度相对低一些，但是它的通用性好，适于各种有紫外吸收的分子检测。荧光检测器灵敏度高，检测下限可达 10^{-15}mol 级，用激光诱导的荧光检测器灵敏度更高。目前用于 DNA 分析的毛细管是利用激光诱导的荧光检测。另外还有电化学、激光热化学以及质谱检测器等。

5.3.3 毛细管的填充物

毛细管内的填充物有高粘度交联的聚丙烯酰胺，这种凝胶毛细管柱制备困难，寿命短，而且管内空间有限，渐渐被具有同样筛分作用的低粘度的线性聚合物溶液代替，称这种电泳为毛细管无胶筛分电泳。多聚溶液对 CE 的 DNA 分离具有重要作用，多聚物的类型和浓度决定于分辨率。目前 PE 公司生产的 POP4，POP6 就是一种液体的胶液，灌胶方便。POP - 4 和 POP - 6，分别代表 perfomance optimized polymer 线性浓度 4% 和 6% 的非交联的二甲基聚丙烯酰胺。溶液中含有高浓度的尿素，使 DNA 变性。

POP - 4 通常用于 STR 分型，POP - 6 的粘度高，电泳时间延长，分辨率提高，应用于测序。在注射每个新样品前都要更换多聚液，即重新从注射器中泵到毛细管中，这个过程

类似于在样品加到胶前重新自动灌制凝胶板。每次分离样品只需几 μl，一满管注射器的胶液可以电泳分离 100 或更多个样品。

5.3.4 毛细管电泳仪

目前已有商品毛细管电泳仪，例如 ABI 公司产品 Prism 310 基因分析仪，为单道毛细管电泳。配有 GeneScan 和 GeneTyper 两种软件，既能作 DNA 序列测定，也能作 STR 基因座分型。PCR 扩增产物自动进样，结果自动检测记录，分辨率达到一个碱基的差异，分型准确，操作简单，已成为常规检测仪器。该仪器改进型 Prism 310，安装有 16 个毛细电泳道，适用于大规模群体调查。

5.4 斑点杂交

斑点杂交（dot blot）是检测 DNA 片段的另一种方法，利用与待测序列互补的寡核苷酸探针杂交检测样品的序列。将 DNA 变性后直接点样于硝酸纤维膜或尼龙膜上，或通过点样器加样，称为斑点印迹法（图 5－7），常用于基因组中序列多态性检测及定量研究，对于那些长度相同但序列不同的 DNA 片段，电泳方法很难将它们分离，可以利用杂交方法将它们分离、检测。其缺点是不能鉴定所测基因的分子量。根据杂交点的形状分为斑点印迹法与狭缝印迹法（slot blot）。斑点印迹为圆形，而狭缝印迹为线状（图 5－8）。一般说来，前者更为清晰，定量更为准确。随着斑点技术推广应用，也可将序列特异的探针先点在膜上，用待测的 DNA 片段带上标记，进行杂交分析，此种杂交被称为“反向杂交”。反向杂交的优点是一次杂交可以同时进行多项分析、多个序列分析，简便、快速、省时。斑点杂交膜制备与检测方法同本章 5.1.1（3）部分。

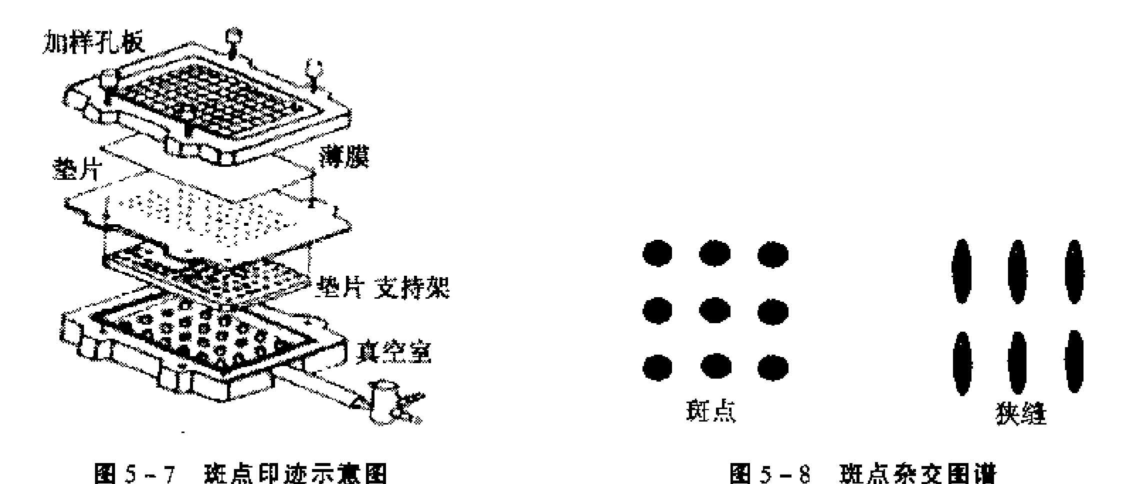


图 5－7 斑点印迹示意图

图 5－8 斑点杂交图谱

第六章 限制性酶切和分子杂交

DNA 限制性酶切和分子杂交技术，不仅应用于限制性内切酶片段长度多态性分析，而且目前还广泛应用于序列多态性检测，已成为法医 DNA 分析常用的基本技术之一。

6.1 限制性内切酶酶切

6.1.1 概念

限制性核酸内切酶（restriction endonucleases，RE）（简称限制酶），是微生物体内一种水解酶，它能识别特异性的 DNA 序列，并在其识别序列内水解磷酸二酯键，将 DNA 切成较小的片段，酶切后的 DNA 片段的 5′端为磷酸基团 P，3′端为羟基 - OH。这类限制酶已成为当代核酸研究中不可缺少的工具酶。

1968 年瑞士巴塞尔生物研究中心 Arber W 实验室首次从大肠杆菌 B 株细胞中部分纯化了Ⅰ型限制性内切酶。1970 年，美国约翰·霍帕金斯大学的 Smith HO 和 Wilcox KW 在世界上第一个发现了Ⅱ型限制性内切酶—Hind Ⅱ。他们用 DNA 黏度测定法证实了流感嗜血杆菌 d（Haemophilus influenzae d）的细胞提取物，可以降解噬菌体 P22 的 DNA，但不能降解其自身的 DNA。Smith 等人提纯了这个 DNA 水解酶，并鉴定了它的识别特异性和切割位点。其特异性和切割位点是：

5′ - GpTpPy ↓ pPupApC - 3′

3′ - CpApPup ↑ PypTpG - 5′

限制性内切酶主要存在于原核细胞中。它们分布很广，有的来自于蓝藻，如 Ava Ⅰ和 Nsp Ⅰ；有的来自于革兰氏阳性菌的嗜血杆菌（如 Hinc Ⅱ，Hind Ⅲ）、葡萄球菌（Sau3A Ⅰ）、革兰氏阴性菌的大肠杆菌（EcoR Ⅰ，EcoR Ⅴ）；还有来自嗜酸热支原体（Thermoplasma acidophilum，Tha Ⅰ）等类微生物。截至 1985 年已从 350 多种微生物中发现了 498 种限制酶，能识别 108 种不同的核苷酸序列。

6.1.2 命名

1973 年 Smith HO 等提出了限制酶的命名原则：（1）以限制酶来源的微生物的学名进行命名，多采用 3 个字母，取其属名的第一个字母（大写）和种名的前两个字母（小写）；（2）若微生物有不同的变种或株系，再加上变种或株系的第一个字母，需大写；同一种菌中发现多种限制酶，则依照发现和分离的前后顺序用罗马数字表示。例如从大肠杆菌（Ecoli）R 株提取的第一种限制酶，命名为 EcoR Ⅰ。Eco 为大肠杆菌 Escherichia coli 菌名的前 3 个字母，R 为其菌株名（RY），Ⅰ表示此酶是从大肠杆菌 R 株中提取到的第一个酶。

6.1.3 限制性内切酶的分类

限制性内切酶有数百种，它们有着共同的特性：只降解双链 DNA 分子，不切单链 DNA 分子；每种酶有其特定的核苷酸序列识别特异性；酶的活性需要二价金属 Mg^{2+} 激活。但是不同的酶之间存在较大的差异，这些差异主要表现在：有些酶不仅要 Mg^{2+}，还需要

其他辅因子的激活；切割位点和识别序列间的距离不同；有的内切酶同时具有甲基化活性。根据限制酶的识别序列、切割特征、催化条件、以及是否有修饰酶活性等这些差异，可将内切酶分为Ⅰ、Ⅱ、Ⅲ型三大类（表 6-1）。通常所指的限制性内切酶是Ⅱ型酶。它们能识别双链 DNA 的特殊序列并在该序列内进行切割，能产生特异性的 DNA 片段。因此Ⅱ型酶是基因工程中有用的 DNA 剪切酶。

表 6-1 三类限制性内切酶的特性

	内切酶的类型		
	Ⅰ	Ⅱ	Ⅲ
DNA 底物	双链 DNA	双链 DNA	双链 DNA
辅因子	Mg^{2+}, ATP, SAM	Mg^{2+}	Mg^{2+}, ATP
识别序列	特异	特异	特异
切割位点	非特定（于识别序列前后 100～1 000bp 范围内）	特定（位于识别序列之中或近处）	特定（切点在识别序列之间 25～27 bp 位置上）
与甲基化作用的关系	同时具有甲基化作用	不具有甲基化作用	同时具有甲基化作用

SAM：S-腺苷酰甲硫氨酸，为甲基供体

6.1.4 Ⅱ型限制性内切酶

Ⅱ型内切酶的识别序列一般为 4～6 个碱基序列，具有高度的专一性。一种限制性内切性酶只能在 DNA 分子中特定的碱基序列（识别座位）处发生作用，将其酶解切开。限制性内切酶的识别序列具有双轴对称的结构，序列转 180°不变，即回文顺序（palindromic sequence），例如 EcoR Ⅰ识别的特异序列：

```
      GAA ⋮ TTC
   ·········⋮·········· 横轴
      CTT ⋮ AAG
          纵轴
```

限制性内切酶切后的 DNA 一般有两种不同类型末端：一种称平齐末端，即在一个水平上切断 DNA 的两条核苷酸链，如 Hae Ⅲ；另一种叫黏性末端，即两条链上的切口错开 2～4 个核苷酸，切下的两条链的末端碱基互补，在 DNA 连接酶作用下可以氢键相连，如 EcoR Ⅰ。

```
5′…G－ C－↓C－C…3′ ──HaeⅢ──→ 5′…G－G＋C－C…3′
3′… C－C－G－G…5′             3′…C－C    G－G…5′

5′…G－↓A－A－T－T－C…3′ ──EcoR Ⅰ──→ 5′…G＋A－A－T－T－C…3′
3′…C－T－T－A－A－G…5′              3′…C－T－T－A－A＋G…5′
```

在双链 DNA 中凡出现这种核苷酸序列，内切酶就可以在箭头所指处将此序列切开。表 6-2 中为常用的限制酶识别序列。

表 6-2　几种限制酶的识别序列

限制酶	识别序列	限制酶	识别序列
Alu Ⅰ	AG↓CT	Kpna	GGTAC↓C
Ava Ⅱ	G↓G（AT）CC	Mbo Ⅰ	↓GATC
Bal Ⅰ	TGG↓CCA	Msp Ⅰ	C↓CGG
Bam Ⅰ	G↓GATCC	Mst Ⅰ	TGC↓GCA
EcoR Ⅰ	G↓AATTC	Pst Ⅰ	CTCGA↓G
	CTTAA↑G		
Hae Ⅲ	GG↓CC	Pvu Ⅰ	CGAT↓CG
Hind Ⅲ	A↓AGCTT	Sau3A	↓GATC
Hinf Ⅰ	G↓ANTC	Taq Ⅰ	T↓CGA
Hpa Ⅰ	GTT↓AAC	Xba Ⅰ	T↓CTAGA

在基因组 DNA 中，4 核苷酸识别序列出现的概率要高于 5、6 核苷酸，对于随机排列的 DNA 链，4 核苷酸序列出现的频率为 $1/4^4$，即平均每 256bp 中就出现一个识别序列，而 5、6 核苷酸序列的频率则分别为 1/1024 和 1/4096。识别序列短则酶特异性低，DNA 被切割成较短的片段。6 核苷酸识别特异性的限制酶在 DNA 中切点少，所切割的片段数少而片段长，特异性较高。由于人基因组 DNA 分子极大，对任何一种限制酶而言均存在许多识别位点。选择何种限制酶是决定 DNA 指纹图特异性的重要条件，使用同一种限制酶，不同个体基因组 DNA 的酶切位置和数目不同，切出片段的长短、数目有个体差异。对同一个体而言，不同限制酶的切割点位置与数目不相同，切出片段的长短、数目也有差异。在美国法庭 RFLP 分析中最常用的酶是 HaeⅢ，欧洲大多数实验室用 Hinf Ⅰ，在我国也以 Hinf Ⅰ为主。

Ⅱ型限制酶不具备甲基化修饰作用，对已甲基化的酶识别序列不能水解切割。不同限制酶对甲基化敏感程度有差异。如 Msp Ⅰ 和 HpaⅡ识别序列都是 CCGG，当此序列甲基化为 C^mCGG 时，Msp Ⅰ仍可切割，而 HpaⅡ则不能酶切该识别点。在人基因组 DNA 中，甲基化仅出现在胞嘧啶，而且常有组织特异性。例如精细胞内小卫星 DNA 中胞嘧啶常以甲基化形式存在，如果选择对甲基化敏感的限制酶，同一个体的精液材料和血液材料 DNA 的酶切可能会出现差异，使 DNA 图谱不一致。因此，对人体材料作 DNA 分型进行同一性认定时，必须考虑选用耐甲基化类型的限制酶。

对于 DNA 指纹图或 DNA 纹印图，选择限制酶有 3 个基本条件：①限制酶识别点应位于 VNTR 位点的旁侧序列，要求尽可能靠近串联重复单位的 5′端和 3′端；②酶特异性和酶活性稳定，不易受酶切反应条件变化的影响；③不受基因组 DNA 中甲基化修饰的影响。在法医学 DNA 纹印图和 DNA 指纹图中常用的酶为 HinfⅠ、HaeⅢ和 PstⅠ。HinfⅠ对 DNA 甲基化敏感，遇到 DNA 甲基化后，不能切割，而 HaeⅢ 对 DNA 甲基化不敏感，遇到 DNA 甲基化仍能切割。

6.1.5 酶活性单位定义

酶活性单位的定义是指在 50μl 反应缓冲液中，在最适反应条件和温度下，1h 完全消化或酶解 1μg λDNA 所需的酶量即为一个酶活性单位。

在实际中，为了使 DNA 完全消化，在 DNA 指纹技术中消化每 1μg 基因组 DNA 实际用量为 3～5u。

6.1.6 酶反应条件

酶反应的条件主要有酶反应缓冲液（包括缓冲体系，pH 值及离子强度）与反应温度。每一种酶拥有自己最适的酶反应条件，不过各种酶之间有许多相似之处。反应体系中的 pH 值一般为 7.4～8.0，以 Tris－HCl 为缓冲体系。反应需要 Mg^{2+} 的激活，并且大部分的酶在反应时需要有巯基化合物（2－巯基乙醇或者二硫苏糖醇）的存在。许多酶还需要 NaCl，不同种的酶对 NaCl 的需要量各有差别，表 6－3 是部分内切酶的反应条件。

表 6－3　限制性内切酶的反应条件

酶	pH	Tris－HCl (mmol/L)	NaCl (mmol/L)	$MgCl_2$ (mmol/L)	2－巯基乙醇 (mmol/L) /Triton X－100 (%)	温度（℃）
AccI	7.5	10	50 (K^+)	6	6	37
Bam HI	7.9	6	150	6	0	37
Eco RI	7.5	100	50	5	0	37
Hae III	7.4	6	50	6	6	37
Hind III	8.0	50	50	10	0	37
Hinf I	7.4	6	100	6	6	37
Hpa I	7.4	10	20 (K^+)	10	1.0 (DTT)	37
Hpa II	7.4	10	6 (K^+)	10	1.0 (DTT)	37
Kpn I	7.5	6	6	6	6	37
Msp I	7.4	10	6 (K^+)	10	1.0 (DTT)	37
Mst I	7.4	6	150	6	6	37
Nae I	8.0	6	20	6	6	37
Pst I	7.5	10	100	10	0	37
Rsa I	8.0	6	0	6	6	37
Sau3A I	7.5	6	50	5	0	37
Taq I	8.4	10	100	6	6	65

大多数酶的反应在 37℃下进行。但是，不少酶的最适反应温度不是 37℃。例如Sma Ⅰ为 25℃、Apa Ⅰ 30℃。有些限制酶的最适反应温度则高于 37℃，例如 Mae Ⅰ 为 45℃，还有些限制性内切酶的最适反应温度可高达 60℃，例如 BstE Ⅰ 是 60℃，Taq Ⅰ 是 65℃。消化反应的温度低于或高于最适温度，都会影响核酸限制性内切酶的活性，甚至最终导致完全失活。

6.1.7 双酶解反应条件

当选用的两种酶所需反应条件相同时，可以在同一反应体系中加入两种酶同时酶解 DNA。比较两种酶的反应体系，如果它们的 pH 值、Tris－HCl 和 $MgCl_2$ 的浓度仅稍有差别，

也可以用同一缓冲液，加入两种酶同时进行酶解反应。当反应体系中两种酶所需的 NaCl 浓度或 pH 值等有较大差异时，就要考虑分开进行两种酶的酶解反应。几种做法是：

一是第一个酶反应后，调整酶反应体系中的离子强度和组分，再进行另一个酶的反应。例如，Sma Ⅰ的缓冲体系为 6mmol/L Tris－HCl（pH8.0～8.5），20mmol/L KCl，6mmol/L $MgCl_2$，6mmol/L 2－巯基乙醇，与常用的核心缓冲液不同，因此一般不能与其他限制性内切酶合用一种缓冲液同时酶解。当 SmaⅠ与其他酶进行双酶解时，可以先用 SmaⅠ进行反应，再用 HCl 调 pH 至 7.5，并补充适量的 NaCl 后，再加上其他的内切酶，如 XbaⅠ，EcoRⅠ等进行酶解反应。

二是先用一种内切酶进行酶解，然后用乙醇沉淀酶解产物，得到的沉淀溶解于第二种酶的反应缓冲液中，再加酶进行酶解。

三是第一种酶解之后，进行琼脂糖凝胶电泳，再从凝胶中回收出各酶切片段，进行第二种酶的消化。

6.1.8 酶解反应终止方法

（1）加 EDTA 以螯合 Mg^{2+}，使酶失去辅因子而停止酶反应；

（2）65℃加热 5～10min。但某些酶在高温下仍有活性，不能用此方法终止这些酶反应。表 6－4 列出了部分限制性内切酶加热灭活的情况。

表 6－4　限制性内切酶的加热灭活

酶	热灭活的敏感度	酶	热灭活的敏感度
Alu I	+	Kpn I	+
BamH I	−	Mbo I	+
EcoR I	+/−	Msp I	+
EcoR II	+	Nco I	+
Hae III	−	Pst I	+/−
Hind III	−	Sau3A I	+
Hinf I	−	Taq I	−

"+"表示65℃加热有95%以上失活；"−"表示酶未被加热灭活；

"+/−"约95%酶失活，但仍能看到部分酶解产物。

（3）加 SDS 到 0.1%的终浓度或加尿素至 0.5mol/L，使酶蛋白解聚变性；

（4）用等体积的酚抽提酶解产物。这种方法灭酶活性最彻底，灭活后的样品用乙醇沉淀法回收。

据实验的具体需要，可以选取合适的终止方法。最常用的终止反应液的组分是：50%（v/v）甘油，100mmol/L EDTA－Na_2（pH8.0），1%SDS，0.02%溴酚蓝。

6.1.9 影响限制性内切酶活性的因素

1.DNA 的纯度

限制性内切酶消化底物 DNA 的反应效率在很大程度上取决于被消化 DNA 本身的纯度。DNA 样品不纯，酶解反应不完全，甚至不被酶解。

（1）DNA 制剂常含有 RNA，虽然它的存在不影响酶的反应速度，但 RNA 可和酶蛋白

发生非特异性结合后可减少酶的有效浓度，使酶解不彻底；另外，RNA 在凝胶电泳所呈现的带会掩盖在该带范围内 DNA 片段的显现，干扰片段的检测。

（2）一般说来，少量的蛋白质的污染对 DNA 的酶解影响不大，但如果有来自其它核酸酶的蛋白质，就会干扰酶切反应和影响酶解产物。结合蛋白（binding protein）能和 DNA 发生非特异的结合，封闭 DNA 的识别序列，影响酶切，使 DNA 和蛋白质的复合物在凝胶电泳上迁移较慢，改变了 DNA 片段的电泳轨迹。

（3）DNA 制剂中的其他杂质，例如酚、氯仿、乙醇、EDTA、SDS、NaCl 等，也会影响酶的反应速度，以至改变识别特异性，出现酶的第二活性。

2. DNA 浓度的影响

在 DNA 的限制性内切酶降解反应中，常常是酶量超过底物量，这样就不用考虑底物浓度对反应的影响。

3. DNA 的甲基化

部分Ⅱ类内切酶不具有甲基化功能，识别序列中特定核苷酸的甲基化会明显影响酶的活性。有些限制性内切酶不能切割甲基化的核苷酸序列。例如 Hinf Ⅰ识别序列 – G↓ANTC – 中胞嘧啶甲基化，成为 – G↓ANTCC^m – 后，Hinf Ⅰ则不能识别此序列，因此不能从此处切割 DNA。

4. 酶消化反应的温度

DNA 消化反应的温度是影响限制性内切酶活性的另一个重要因素。不同的限制性酶，具有不同的最适反应温度。相当一部分酶，在反应温度升高后活力不稳定，像 EcoR Ⅰ在 42℃时就失活，因此可以采用 65℃加热来灭活。但有些酶像 BamH Ⅰ，在高于 37℃温度下活力更高。

5. DNA 的分子结构和构型

限制性内切酶能识别双链 DNA 上的特异序列，但同一链上不同部位的识别序列被酶切割的速度并不一致。例如，λDNA 中有 5 个 EcoR Ⅰ的酶切点，其中两个切点的反应速率相差 10 倍。

DNA 分子的不同构型对核酸限制内切酶的活性也有很大的影响。某些限制性内切酶切割超螺旋的质粒 DNA 或病毒 DNA 所需要的酶量要比消化线性 DNA 的高，甚至可高达 20 倍。

6. 限制性内切酶的缓冲液

限制性内切酶的标准缓冲液的组分包括 $MgCl_2$、NaCl 或 KCl、Tris – HCl、二硫苏糖醇（DTT）以及牛血清白蛋白（BSA）等，这些组分的变化会引起限制性内切酶活性的变化与识别序列特异性的改变。

（1）阳离子　酶活性的正常发挥依赖二价的阳离子，通常是 Mg^{2+}。浓度在 5mmol/L 以上，但在 5～30mmol/L 浓度范围，未见明显的活力变化。反应中不能加入 EDTA，因它能和 Mg^{2+} 结合，降低游离的 Mg^{2+} 浓度，使酶活力降低。反应体系中常含有 NaCl，以提供一定的离子强度。不同的酶要求 NaCl 浓度不同。不适当的 NaCl，会降低限制酶的活性，还可能导致识别序列特异性的改变。

（2）pH 值　绝大多数限制酶在 pH7.4 的条件下，其活性最佳。最常用的缓冲液是 Tris

-HCl。

(3) 巯基试剂　巯基试剂是保持某些内切酶的稳定性的重要因素之一，同时能降低污染杂质对酶活性的干扰。

(4) 在酶制剂中常加入 BSA 或白明胶来稳定酶活性。加入 BSA，可以减少蛋白水解酶对酶蛋白的水解，也能减轻非特异性吸附。常用的 BSA 量为 100μg/ml，BSA 制剂不能混有核酸酶活性。配制的溶液需要经过滤膜过滤消毒。一般地，加入消毒过的无核酸酶活性的 BSA 对酶反应无干扰，但 BSA 过多时会和 DNA 结合，使电泳轨迹不清晰。

6.1.10 限制性内切酶的星号活性

有些限制性内切酶的切割特异性受所用缓冲液成分的影响比较明显。例如高浓度的限制性内切酶、高浓度的甘油（75% v/v）、Mn^{2+} 取代 Mg^{2+}、残留乙醇以及高 pH 值等情况下，有些内切酶识别序列的特异性便会发生变动，在其识别序列以外的其它位置切割 DNA 分子，通常叫做星号活性（star activity）或叫第二活力（secondary activity）。例如，正常的情况下，EcoR Ⅰ在识别序列 GAATTC 处切割，当缓冲液中的甘油浓度超过 50%（V/V），其识别序列的特异性就会发生变动，可在 AATT 序列处切割 DNA。容易出现星号活性的酶有 Taq Ⅰ、EcoR Ⅰ、Hind Ⅲ、Kpn Ⅰ、Pst Ⅰ、Hinf Ⅰ及 EcoR Ⅴ等。

理想的酶解作用是保证 DNA 指纹技术成功的重要条件。如何得到理想的酶解反应，可以归纳为：

(1) DNA 的纯度要高，不能含有 RNA 和蛋白质，也不能含有过高的 EDTA，更不能含有酚、乙醇、SDS 和其它有机溶剂。

(2) 采用推荐的酶反应最适条件。

(3) 选用高质量的酶制剂，使之不含非特异性核酸酶、外切酶和磷酸酶等杂酶活性。

(4) 采用适宜的酶/DNA 比例：2～10u/μg，避免过高的酶浓度。

(5) 采用适宜的反应温度和反应时间，尽量不要进行长时间的酶解反应。

6.1.11 限制性内切酶消化反应操作注意事项

(1) 实验中所用的 Eppendorf 离心管和吸头要干净，DNA 酶、酚和去污剂的污染将影响酶切效果。

(2) 大多数厂家都以浓缩形式提供限制酶，不同的生产厂家往往使用不同的反应条件，按说明书进行操作。

(3) 不同限制酶的缓冲液主要差别在 NaCl 的浓度。当用两种或两种以上限制酶切割 DNA 时，如果酶所要求的缓冲液条件相同，则两种酶可同时切割。如果这些酶所要求的缓冲液不同，则可采用以下两种方法：①先用在低离子强度的缓冲液中活性最高的酶切割 DNA，然后加入适量 NaCl 及第二种酶，继续反应；②用第一种酶消化 DNA 后，以酚-氯仿进行抽提，再经乙醇沉淀，DNA 重新溶解再进行第二种酶酶切。

(4) 酶切时，使酶在冰箱外放置的时间应尽可能缩短，从冰箱内取出贮酶管，立即放置于冰上，每次取酶时应换一个无菌吸头。

(5) 尽量减少反应中的加水量以使反应体积减到最小，但要确保酶体积不超过反应总体积的 1/10，否则，酶的活性将受到甘油的抑制。

6.1.12 **内切酶消化产物酶切程度检测**

由于物证检材的多样性与复杂性，提取的 DNA 中经常有异物污染，如染料、色素等，均可能抑制酶切反应。DNA 样品液中残留的蛋白质、酚、SDS 等也会干扰酶切反应，所以在进行凝胶电泳分离前应对内切酶消化 DNA 是否完全进行检测，只有确认 DNA 酶切达到预想要求时才可以继续进行下一步操作。

酶切效果检验方法：配制 0.6%琼脂糖凝胶（6 cm× 8cm），以 1×TAE 为缓冲液。取少量酶切产物（3μl），加样在凝胶中，在相邻泳道内加入 DNA 分子量标准和没有酶解的 DNA 样品作为参考，60V 电压下电泳大约 1h，凝胶在 0.5μg/ml EB 溶液中染色 0.5h，紫外灯下观察酶切结果。完全消化的 DNA 样品在凝胶中表现为一有各种大小分子组成的荧光强度均匀的柱状（即血涂片样的一条电泳轨迹），加样孔附近没有荧光；完全没有酶解的 DNA 被压缩成一条谱带，靠近凝胶加样孔附近；部分消化 DNA 表现为一条尖而高的类似火箭电泳的图形，加样孔附近有荧光，在自显影照片中表现为一黑色柱子。

6.2 分子杂交

分子杂交基本原理是：具有一定同源性的两条单链 DNA，在一定的 DNA 复性条件下，可按碱基互补原则退火形成双链，根据带标记的已知序列（探针）检测待测 DNA 分子的序列差异。待测靶序列可以是核基因组 DNA，也可以是 PCR 扩增产物。

DNA 分子杂交方法可分为固相液相杂交、液相分子杂交和原位杂交。法医 DNA 分析中常用的是固相液相杂交，亦称膜上印迹杂交，先将待测 DNA 片段电泳分离，然后用印迹技术将分离的 DNA 片段原位转移到特定的固相支持物上或者直接将待测 DNA 直接点在支持物上，再用标记的 DNA 探针与之杂交。洗去未杂交的游离的探针分子，通过放射自显影等方法显示标记的探针位置，检出靶 DNA 片段的位置。

6.2.1 DNA 转移

6.2.1.1 固相支持物的选择

凝胶中的 DNA 分子不适合直接进行杂交，一是因为琼脂糖凝胶机械强度差，易碎，不易操作；二是这种胶上直接杂交后背景高，因此，需要转移到一理想的固相支持物上。性能良好的固相支持物应具有以下几个条件：

1. 具有较强结合核酸分子的能力，结合稳定牢固，在杂交、洗膜过程中核酸分子不会脱落；
2. DNA 分子被吸附后应不影响其与探针分子的杂交反应；
3. 对探针分子的非特异性吸附少，杂交信号背景低；
4. 具有一定的机械强度，韧性好，便于操作。

固相支持物的种类较多，各种材料的性质不同，各有特点，适合不同的要求（表 6－5）。较为常用的是硝酸纤维素膜和尼龙膜。

表 6-5　用于固定 DNA 的各种材料性质

	硝酸纤维素膜	尼龙膜	带正电荷的尼龙膜	活化纸
应用	ssDNA,RNA 蛋白质	ssDNA,dsDNA, RNA,蛋白质	ssDNA,dsDNA,RNA, 蛋白质	ssDNA,RNA
结合能力 (μg 核酸/cm^2)	80~100	400~600	400~600	2~40
抗张强度	差	好	好	好
核酸结合形式	非共价	共价	共价	共价
结合核酸最小分子量	500 核苷酸	500 核苷酸	50 核苷酸	5 核苷酸
重复杂交性能	差,信号易丢失	好	好	好

1. 硝酸纤维素膜

硝酸纤维素膜（nitrocellulose membrane）具有较强的吸附单链 DNA 和 RNA 的能力，特别是在高盐浓度下，结合核酸能力可达 80~100μg/cm^2。吸附的单链 DNA 或 RNA 经真空烘烤后，依靠疏水性非共价作用而结合在硝酸纤维素膜上。硝酸纤维素膜非特异性地吸附蛋白质较弱、具有杂交信号本底低等优点，应用于 Southern、Northern、斑点印迹及克隆筛选中。

硝酸纤维素膜虽是应用较早、最广泛的一种固相支持物，但在 DNA 分子杂交应用中有一定局限性。原因是：①硝酸纤维素膜结合 DNA 不十分牢固，在杂交及洗膜的过程中 DNA 容易脱离，尤其在高温情况下，杂交效率明显下降，另外膜不能进行反复杂交。②硝酸纤维素膜质地较脆，经烘烤后易碎，操作需特别小心。③硝酸纤维素膜与核酸的结合要求在 10×SSC 的高盐浓度条件下，低盐浓度时结合 DNA 效果不佳，因此不适宜于电转印迹法。④硝酸纤维素膜对于小分子量 DNA 片段，特别是 <200bp 以下的 DNA 片段，结合能力不强。

2. 尼龙膜

尼龙膜（nylon membrane）是目前较理想的一种核酸固相支持物，它有多种类型。其中经过修饰带有正电荷的尼龙膜结合核酸能力更强。如 Amersham 公司产品 Hybond-N 是普通尼龙膜，Hybond-N+、Bio-Rad 公司的 Zeta-Probe、DuPont 公司的 Gene-Screen Plus 及 Pharmacia 公司的 GeneBind 是正电荷修饰的尼龙膜。

尼龙膜结合单链及双链 DNA 和 RNA 的能力比硝酸纤维素膜更强，可达 350~500μg/cm^2。经烘烤或紫外线照射或碱处理后，DNA 分子可牢固地结合在尼龙膜上，特别是用短波紫外线照射后，核酸中部分嘧啶碱基可与膜上的带正电荷的氨基相互交联，使结合更加牢固。尼龙膜可重复杂交，一次杂交后，探针分子经碱处理被洗脱下来，然后可与第二探针杂交。对于小分子量的核酸片段亦有较强的结合能力；在低离子强度条件下也可较好地结合 DNA，因此也比较适合于电转印迹法。另外，它的韧性较强，操作方便，是目前最为常用的固相支持膜。

尼龙膜的缺点是杂交信号本底较高，需在预杂交液中加非特异性封闭试剂，以克服本底高的问题。

6.2.1.2 印迹方法

1. 萨森转移（Southern blot）

萨森转移是 E.Southern 于 1975 年创立的印迹转移方法。它是利用毛细管虹吸作用由转移缓冲液带动 DNA 分子转移到固相支持物上，将电泳分离的 DNA 片段从凝胶转移到硝酸纤维素膜或尼龙膜上的过程。

（1）取下电泳后的凝胶，将凝胶浸泡于适量的碱变性液（1.5mol/L NaCl，0.5mol/L NaOH）中，置室温 1h。

（2）将凝胶用去离子水漂洗 1 次，然后浸泡于适量的中和液（1mol/L Tris－HCl，1.5mol/L NaCl，pH8.0）中 30min，不间断地轻轻摇动。更换新鲜中和液，继续浸泡 15min。

（3）如图 6－1 所示，在一塑料或玻璃平台上铺一层 Whatman 3MM 滤纸，此平台要求比凝胶稍大。将此平台置于一搪瓷盘或玻璃缸中，搪瓷盘中盛满 20 × SSC（3mol/L NaCl、0.3mol/L 柠檬酸钠，pH 7.0）或 10 × SSC（也可用 10 ×或 20 × SSPE 代替），滤纸的两端要完全浸没在溶液中。将滤纸用 20×SSC 湿润，用一玻璃棒将滤纸推平，排除滤纸与玻璃板之间的气泡。

（4）裁剪下一块与凝胶大小相同或稍大的尼龙膜。注意操作时要戴手套，不可用手触摸。将尼龙膜漂浮在去离子水中，使其从底部开始向上完全湿润。然后置于 2 × SSC 中至少 5min，如果滤膜不能被湿润则不能用。

（5）将中和后的凝胶，置于上述铺了一层厚滤纸的平台中央，两者之间不要有气泡。在凝胶的四周用封口膜封严，以防止在转移过程中转移液直接从容器中流向吸水纸而不经过凝胶，从而使转移效率降低。

（6）将湿润的滤膜小心覆盖在凝胶上，膜的一端与凝胶的加样孔对齐。排除两者之间的气泡。注意膜一经与凝胶接触即不可再移动，因为从接触的一刻起，DNA 已开始转移。

（7）将两张预先用 2 × SSC 湿润过的与滤膜大小相同的滤纸覆盖在滤膜上，排除气泡。

（8）裁剪一些与滤膜大小相同或稍小的吸水纸（约 5～8cm 厚），放在滤纸上。在吸水纸之上放一玻璃板，再压一重约 500g 的物品（图 6－1）。转移液将在吸水纸的虹吸作用下从容器中转移到吸水纸中，从而带动 DNA 从凝胶中转移到滤膜上。转移 8～24h，其间换吸水纸 1～2 次。

（9）弃去吸水纸和滤纸，将凝胶和滤膜置于一张干燥的滤纸上。用软铅笔标明加样孔的位置。滤膜浸泡在 6×SSC 溶液中 5min 以去除琼脂糖碎块。此时凝胶用 EB 染色，在紫外灯下可检查转移的效率。

（10）用滤纸吸干膜，然后置于两层干燥的滤纸中，真空下 80℃烘烤 2h，使 DNA 固定于膜上。因为转移后的膜与 DNA 结合并不十分牢固。

此尼龙膜即可用于下一步的杂交反应。如果不马上使用，可用保鲜膜包好，室温下置真空中保存备用。如果时间较长，应置于－20℃。

2. 真空转移（vacuum transfer）

经典的利用毛细管虹吸作用的转移法效率不高，特别是对于分子量较大的 DNA 片段，

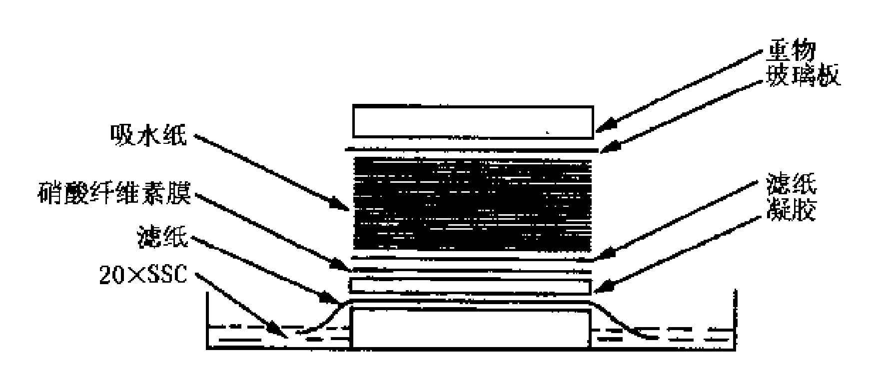


图 6-1　Southern 印迹转移

需要长时间转移。目前大多数实验室采用真空转移方法，其原理是利用真空作用将膜缓冲液从上层通过凝胶抽到下层真空室中，同时带动核酸片段转移到凝胶下面的尼龙膜或硝酸纤维膜上（图 6-2）。真空转移法是一种简单、快速、高效的 DNA 和 RNA 印迹法。

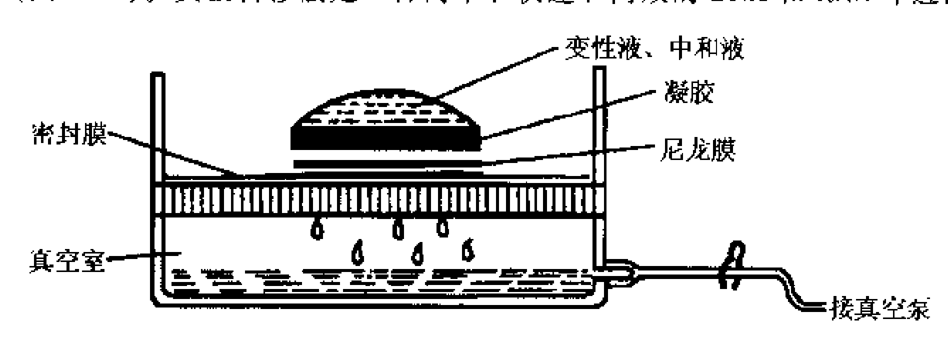


图 6-2　真空转移示意图

真空转移法的最大优点是快速，可在转膜的同时进行 DNA 的变性与中和，整个过程只需 30min 至 1h。具体操作过程在 DNA 指纹章节介绍。

3. 电转移法

电转法（Electrophoretic transfer）：利用电场的电泳作用将凝胶中的 DNA 转移到固相支持物上，是近年来发展起来的一种简单、迅速、高效的 DNA 转移法（图 6-3）。一般只需 2～3h 即可完成转移过程。特别是对于用虹吸法不理想的大片段 DNA 的转移较为适宜。

4. 斑点印迹

斑点印迹法（dot blot）与狭缝印迹法（slot blot）是由 Southern 印迹衍生而来的两种类似的印迹方法。将 RNA 或 DNA 变性后直接点样于硝酸纤维膜或尼龙膜上，或通过点样器加样，称为斑点印迹法，常用于基因组中特定基因及其表达的定性及定量研究，也是实验中常用的技术之一。与 Southern 印迹法相比，其优点是简单、快速，可在同一张膜上同时进行多

海绵　电极　电　凝胶支持夹　海绵　滤纸　凝胶　尼龙膜　滤纸　凝胶支持夹　海绵

图 6-3　电转移示意图

个样品的检测。

6.2.2 DNA 探针与标记

6.2.2.1 DNA 探针

1. 探针选择的原则

DNA 探针是一段已知序列的 DNA，带上标记后用来检测靶序列。探针选择原则为：(1) 高度特异性；(2) 杂交稳定性，重复性好；(3) 来源容易，制备简便；(4) 靶基因座多态性高；(5) 突变率低。应用探针前，必须对筛选的探针进行全面评估，包括所测基因座的染色体定位、多态性片段的长度范围、等位基因数及在群体中的频率分布，基因座的突变率等。实际选择探针最基本的原则是应该具有高度特异性，兼而考虑来源是否方便等其他因素。

探针由不同实验室筛选并命名，因此探针的命名至今尚未统一。

2. DNA 探针种类

DNA 探针种类很多，按探针来源及性质分类，DNA 探针有基因组探针、人工合成寡核苷酸探针及 cDNA 探针。cDNA 探针不太用于法医学中，基因组探针与合成寡核苷酸优缺点比较如表 6-6。早期的探针均是基因组探针，采用分子克隆技术和质粒培养等方法制备探针。近年来 PCR 技术的应用，PCR 也可以制备基因组 DNA 探针，使探针制备更加方便。不过，随着标记技术的发展，越来越多地采用合成的寡核苷酸探针。

表 6-6 几种 DNA 探针比较

探针	优点	缺点
双链 DNA 探针	不需再一次克隆	探针需要变性
	能选用各种标记方法	杂交反应中可能出现自身复性
	杂交适宜的温度范围较宽	杂交体不如 RNA 探针稳定
单链 DNA 探针	探针不含载体序列	需要再克隆到 M13 载体中
	不需变性过程	杂交体不如 RNA 探针稳定
	杂交时无自身复性	技术较困难
合成寡核苷酸探针	不需要克隆	标记方法受到限制
	无自身杂交	携带标记分子少
	组织穿透性好	探针设计可能会有错误
	可以根据氨基酸序列制备探针	需要 DNA 自动合成仪合成
	能避免有关序列之间的同源性	杂交体不如 RNA 探针稳定
	稳定性好	只能制备短序列寡核苷酸探针

6.2.2.2 DNA 探针制备

1. 基因组 DNA 探针制备 ——利用基因重组技术制备探针

首先选择限制性内切酶，对基因组 DNA 进行剪切，选出某种特定的小卫星 DNA 片段。然后选择合适的基因转移载体如质粒，对其进行限制性内切酶的切割，最后与插入的目的 DNA 片段连接、重组，即可获得含有目的 DNA 片段的重组质粒。将质粒转化到大肠杆菌中，从阳性克隆中筛选得到含质粒的菌株，经培养大量繁殖细菌以扩增质粒，再从细菌中将质粒提取出来，提取纯化质粒 DNA，用限制性内切酶将插入的 DNA 片段切出，制

备所需的 DNA 探针。

2. 寡核苷酸探针制备

(1) 化学合成

基因的合成研究始于20世纪50年代，合成的基本原理是：将所要合成的寡核苷酸链的末端核苷（或核苷酸基）先固定在一个不溶性的高分子表面上，然后再从此末端核苷开始逐个接长寡核苷酸链。每次接成一个核苷或一个片段，则经历一轮循环，接长的链始终被固定在不溶的固相载体（或支持物）上，过量的未反应物或分解物则通过过滤或洗涤除去。当整个链的增长达到所需的长度后，再将寡核苷酸链从固相载体上切落下来脱去保护基后，经过分离纯化得到所需要的最后产物。图 6-4 表示了固相三酯法合成寡核苷酸的简单过程。在固相合成中，由于使用了固相载体，将合成的寡核苷酸固定，而过量的未反应物和副产物只需用洗涤法除去，简化了每次连接产物的分离纯化步骤，加快了合成的速度。

1）末端固定

2）去保护及激活

四唑

3）缩合及偶联

(<2%)　(>98%)

偶联B

I_2

4）氧化

循环

图 6-4 亚磷酸酰胺合成寡核苷酸原理示意图

对于几十个 bp 长的寡核苷酸探针如（CAC)$_5$/（GTG)$_5$ 可以利用化学合成方法制备。目前有 DNA 自动合成仪自动合成，更有商业公司的合成服务，需要制备几十个 bp 长的寡核苷酸探针，只需提供 DNA 探针序列，委托公司合成。

目前，基因组小卫星 DNA 探针制备也改进成用化学合成小卫星核心序列，并且在合成过程中标记上标记物，使 DNA 探针制备、标记一步完成，更加简便。

（2）T4 连接酶法制备 DNA 探针

T4 连接酶具有将两个或多个单链状态的 DNA 片段连接成较长的 DNA 片段的功能，可以利用 T4 连接酶制备数百 bp 长的寡合核苷酸探针。具体做法是：要将连接的寡核苷酸加在 T4 连接酶反应液中，在 37 ℃下反应 16h。反应时间越长，短寡核苷酸越多，连接成的寡核苷酸链越长。

（3）利用 PCR 技术制备 DNA 探针

设计 DNA 探针序列两侧互补的 DNA 序列作为 PCR 引物，利用 PCR 技术将 DNA 探针序列特异地扩增出来（图 6-5），目前 PCR 技术可以扩增出数 kb 长的 DNA 片段，足以满足 DNA 探针的制备。

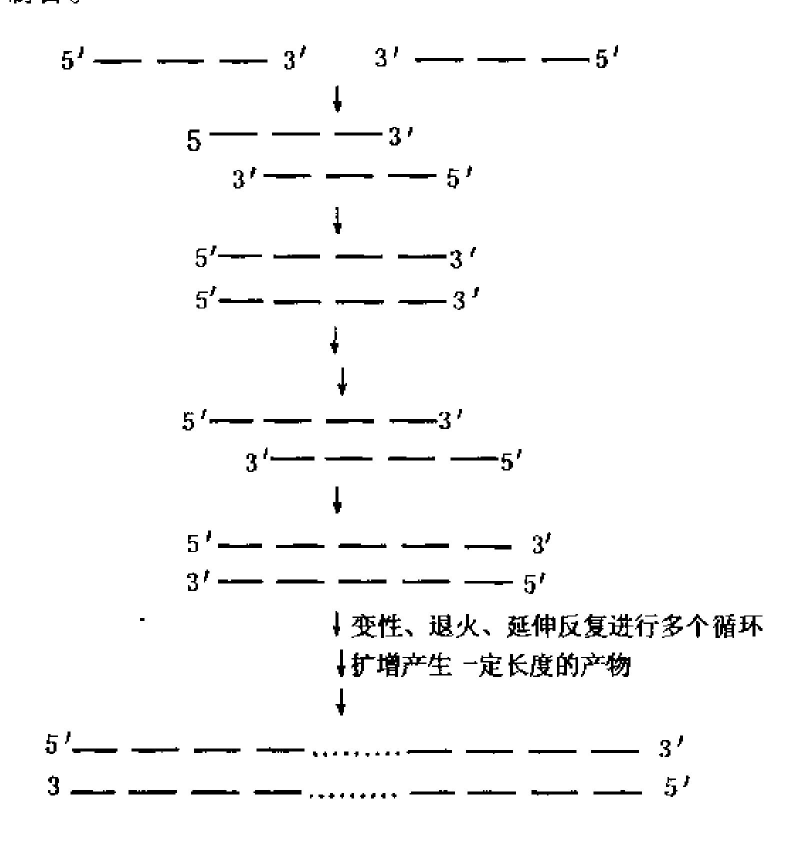


图 6-5 PCR 法制备探针

6.2.2.3 标记物的种类及特性

1. 标记物的选择

DNA 探针只是一段已知序列的 DNA，只有当它带上同位素或其他非同位素等标记物后才能检测显示所要检测的目的 DNA 序列，因此 DNA 探针在进行分子杂交前必须进行标

记。作为一个理想的标记物应该具备以下特点：①高灵敏度；②标记方法简单；③相对稳定，保存时间长；④特异的理化性质如光谱学特征或生化活性；⑤可通过一定的途径与被标记物结合，并且这种结合不改变标记物的基本特性；⑥标记时不影响碱基配对特异性；⑦当用酶促方法标记时，对酶促反应的底物无较大的影响。

探针标记物有同位素和非同位素，前者主要为^{32}P，后者主要有生物素、地高辛（Digoxin）等半抗原、荧光素和酶等（表6-7）。

表6-7　部分标记物

放射性同位素	非放射性标记物			
	生物素	酶类	半抗原	荧光素
^{3}H	生物素	碱性磷酸酶	地高辛	异硫氰酸荧光素
^{32}P	光敏生物素	辣根过氧化物酶	磺基化胞苷	羟基香豆素
^{33}P				
^{35}S				罗丹明
^{125}I				

（1）放射性同位素标记物

放射性同位素衰变时，不稳定同位素变为稳定同位素，同时从原子核内发出射线。发出的射线能感光X片，留下轨迹。如^{32}P和^{35}S等发出β-射线，它是带负电荷的高速离子流，能量很高，穿透能力强。

放射性同位素标记灵敏度高，可以检测到$10^{-14}\sim10^{-18}$克的物质，在最适条件下，可以测出样品中少于1 000个分子的核酸含量。但放射性同位素存在污染环境、有碍健康、废物较难处理和半衰期短等弊端。

（2）非同位素

①生物素：生物素（biotin）属于B族维生素，将生物素连接在核酸探针上，利用亲和素对生物素有极高的亲和力的原理，分子杂交后通过亲和素（Avidin）或链亲和素（Streptavidin）与酶结合，进行化学反应检测。亲和素是一种碱性的四聚体糖蛋白（68KD，pI10.5），每个亚基能结合一个生物素分子。生物素经羟化修饰后可制备成具有各种活性基团的衍生物，成为活化的生物素。活化的生物素能与蛋白、糖类或核酸等偶联，当带有标记物的亲和素与偶联的生物素之间结合后，即可被显示。生物素与亲和素的结合具有专一、迅速和稳定的特点，并能放大检测信号。

②半抗原：在核酸分子上用半抗原标记制备成免疫核酸探针，使之能用免疫学方法检测。地高辛是类固醇半抗原，地高辛通过一个含11个碳原子的连接臂与尿嘧啶核苷酸嘧啶环上的第5组碳原子连接形成地高辛标记的尿嘧啶核苷酸。

③酶：主要有辣根过氧化物酶（horseradish peroxidase，HRP）和碱性磷酸酶（alkaline phosphatase，AP）。辣根过氧化物酶或碱性磷酸酶直接标记在探针上，利用酶的化学反应显色或光化学反应中发出的光进行自显影。

④荧光素标记物有异硫氰酸荧光素（isothiocyanante，FITC），羟基香豆素（resorufin,）

罗丹明（rhodamine），及其衍生物四甲基异硫氰酸罗丹明（tetramethy - rhodamine isothiocyanate，TRITC），氨甲基香豆素醋酸酯（aminomethylcoumarin acetate，AMCA），德克萨斯红（texas red），主要用于原位杂交中。

6.2.2.4 DNA 探针标记

DNA 探针标记方法多种多样，可分为缺口翻译、末端标记、随机引物标记、酶直接标记法及 PCR 反应标记法。标记物分子及标记方法的差异都关系到探针的灵敏度、分辨率和特异性等重要问题。表 6 - 8 为 DNA 探针类型与常用标记方法。目前酶直标法也可用 20bp 以上的合成寡核苷酸。

表 6 - 8 DNA 探针类型与标记方法

探针类型	标记方法
双链 DNA	缺口平移法，随机引物法，酶直标法
单链 DNA	随机引物法，酶直标法
合成寡苷酸	末端标记法，随机引物法

1. 缺口翻译标记法

缺口翻译法（nick translation）或缺口平移法是早期实验室的 DNA 探针标记法。它利用大肠杆菌 DNA 聚合酶Ⅰ（E.coli DNA polymerase Ⅰ）的多种酶促活性将标记的 dNTP 掺入到新形成的 DNA 链，合成带标记物的 DNA 探针。线状、超螺旋及带缺口的环状双链 DNA 均可作为缺口翻译法的模板。首先，极微量 DNase Ⅰ在 Mg^{2+} 的存在下，在 DNA 链上随机形成单链缺口。利用大肠杆菌 DNA 聚合酶Ⅰ的 5′ - 3′核酸外切酶活性在切口处将旧链从 5′ - 末端逐步切除。同时，DNA 聚合酶Ⅰ的 5′→ 3′聚合酶活性的催化下，顺序将 dNTP 连接到切口的 3′ - 末端 - OH 上，以互补 DNA 单链为模板合成新的 DNA 单链。如果在反应液中含有一种或多种标记的核苷酸（如 ^{32}P - dCTP，或 dUTP - 11 - biotin），这些标记的核苷酸将替代原来的核苷酸残基，结果在新合成的一条单链上带有标记物，成为高放射活性的或带生物素的 DNA 探针（图 6 - 6）。

2. 随机引物标记法

随机引物标记法（random primer）是近年来发展起来的一种较理想的 DNA 探针标记方法，已成为实验室 DNA 探针标记的常规方法。利用随机引物，与模板 DNA 结合，在 DNA 酶作用下按 5′→3′方向合成与模板互补的新的 DNA 链。

随机引物（random primer）指含有各种可能排列顺序的寡核苷酸片段的混合物，它可以随机地与核苷酸序列杂交，起到聚合酶反应的引物作用（图 6 - 7）。实验室中所用的随机引物可以通过 DNaseⅠ酶解小牛胸腺 DNA 得到。目前市售的试剂盒中的随机引物是人工合成的片段长度为 6bp 寡核苷酸，含有各种可能的序列（$4^6 = 4096$ 种排列序列）。

将待标记的 DNA 片段变性后与随机引物混合，随机引物结合在 DNA 片段上，在大肠杆菌 DNA 聚合酶Ⅰ大片段（klenow fragment）催化下，合成与探针 DNA 序列互补的 DNA 链。当反应液中含有 $\alpha-^{32}P$ - dNTP 或 dUTP - 11 - biotin 时，即形成放射性或生物素标记的 DNA 探针。

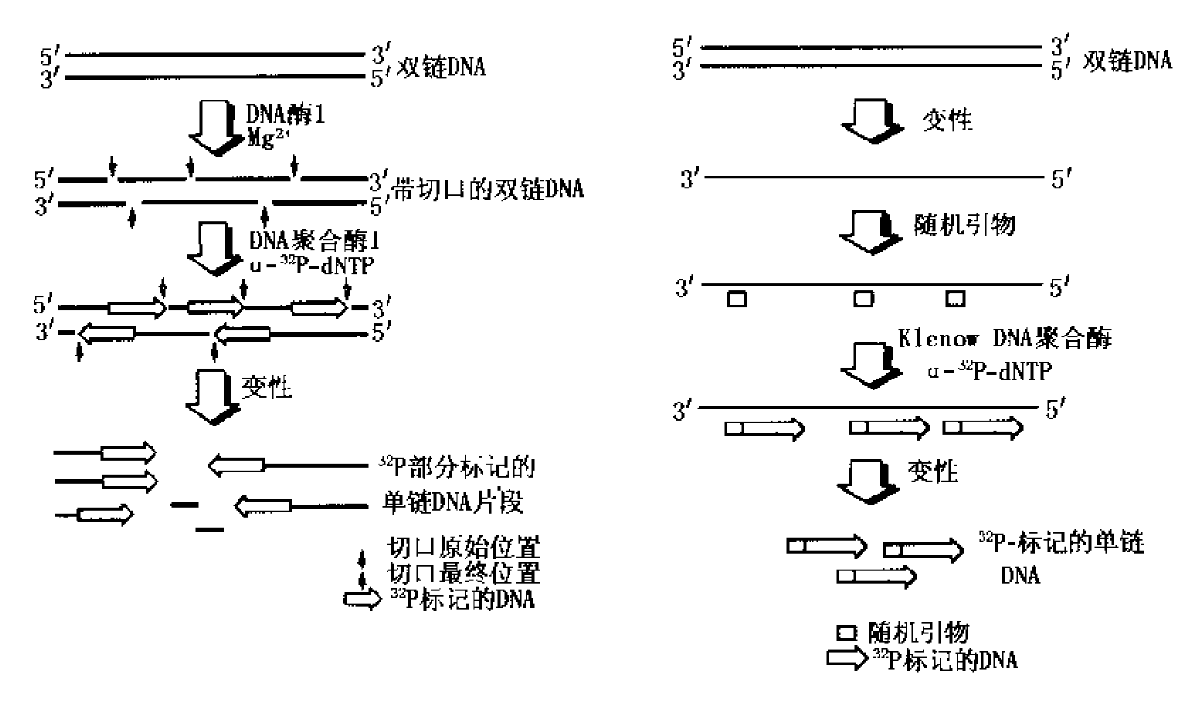


图 6-6　缺口平移法示意图　　图 6-7　随机引物标记法原理示意图

3. 末端标记法（End - Labelling）

与缺口平移法及随机翻译法不同，末端标记法并不将 DNA 片段的全长进行标记，而是只将其一端（5′或 3′端）进行部分标记，将标记物导入线形 DNA 的 5′或 3′端。

5′端标记法需用 T_4 多聚核苷酸激酶（T_4 polynucleotide kinase，PNK），而最常用的标记物是（$\gamma-{}^{32}P$）ATP 和 Biotin - 11 - dUTP。T_4 多聚核苷酸激酶能特异性地将 ^{32}P 从（$\gamma-{}^{32}P$）ATP 转移到 DNA 的 5′ - OH。由于大多数 DNA 的 5′端含有磷酸基团，因此标记前先要用碱性磷酸酶将磷酸基团切除。

DNA 的 3′端可用末端脱氧核苷酰转移酶（terminal deoxytransferase，TDT）来进行标记，该酶能不依赖模板将 dNTP 加到单链或双链 DNA 的 3′末端上，可将单个或多个标记的核苷酸加到 3′端上（图 6-8）。

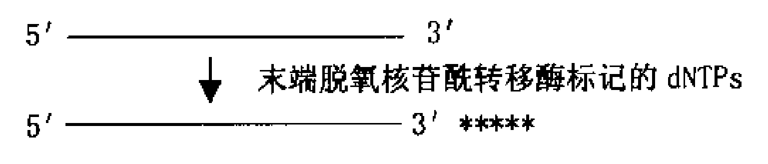


图 6-8　末端转移酶制备 3′末端标记探针原理示意图

4. 酶直接标记法

此方法主要用于非同位素标记物，如辣根过氧化物酶，先与对苯醌（PBQ）结合，再利用对苯醌的活化基团与聚乙烯亚胺（PEI）结合，形成带大量正电荷 HRP - PBQ - PEI 复合物。复合物与带负电荷的 DNA 片段结合，再利用交联剂戊二醛的酰胺键，共价结合交联在一起，形成稳定的酶 - DNA 复合物探针。一个 DNA 探针分子上可以带有多个酶分子

(图 6－9)。

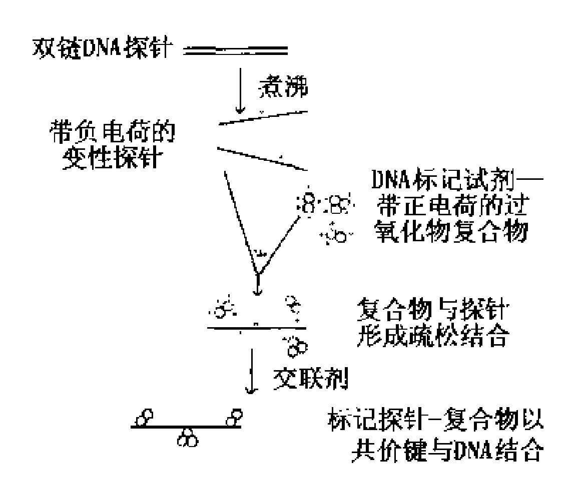


图 6－9 辣根酶标记探针和检测反应原理

5. 合成偶联标记法

在 DNA 合成过程中 5′－末端碱基连接一个带 NH_2的连接物（linker），利用氨基同羟基反应，连接上一个酶分子或荧光分子，直接标记寡核苷酸上，目前许多商业试剂盒如 Cellmark Diagnosis 中使用的探针属于此类方法制成的。

6.2.2.5 DNA 探针的纯化

DNA 探针标记反应结束后，反应液中仍存在未掺入到 DNA 中去的 dNTP 等小分子。如不除去，有时会干扰下一步的杂交反应，因此 DNA 探针标记后需要纯化。纯化的方法有（1）凝胶过滤柱层析法（2）乙醇沉淀法（3）反相柱层析法。

同位素标记的常用 SephadexG－50 凝胶过滤柱层析法。非同位素标记的则常用乙醇沉淀法。

6.2.3 固相液相分子杂交

由斑点或狭缝印迹转移成的滤膜进行的杂交，叫斑点或狭缝杂交。由 Southern 或真空或电转移印迹而来的滤膜杂交叫 Southern 杂交。杂交包括预杂交封闭过程、DNA 探针分子与靶 DNA 片段的杂交和杂交后的洗膜。洗膜的目的是去除在杂交过程中非特异结合的 DNA 探针分子，提高背景信号噪音比，获得理想 DNA 杂交图谱。

1. 杂交体系建立

分子杂交过程，实际上是 DNA 的复性过程。只要待测 DNA 样品中存在与加入 DNA 探针同源或异源的互补序列，在一定条件下，即可退火形成 DNA 双链，即 DNA 杂交体。复性过程相当复杂，并不是个简单变性的一逆反应过程。变性过程可以在一个极短的时间内迅速完成，而复性则需要相对较长的时间。如果使热变性的 DNA 溶液迅速冷却，则只能形成一些不规则的碱基对，而不会完全恢复 DNA 双链结构。但将变性 DNA 溶液置低于 Tm 值 25 ℃ 的温度下维持一段时间后，则可使之回复到天然的双链结构状态。前面已介绍 DNA 复性受许多因素影响，建立杂交体系应考虑到这些因素：

（1）DNA 分子的复杂性：DNA 的复杂性与其分子量成正比，复性速度与 DNA 分子的

复杂性成反比。分子量越大、越复杂，复性速度越慢。如大肠杆菌染色体 DNA 的分子量比噬菌体 φx174 DNA 大 1 000 倍，其复性速度也慢 1 000 倍。

（2）探针 DNA 的片段大小：复性过程是两条单链间互相碰撞过程。碰撞的机会越多，复性的速度越快。大的 DNA 片段在溶液中的扩散速度较慢，碰撞互补单链的机会少，复性慢。小的 DNA 片段在溶液中的扩散速度快，增加了相互之间的碰撞机会。因此在杂交时一般制备成小片段的探针。

（3）DNA 的浓度：浓度越高，单链之间互相碰撞的机会越多，复性越快，因此在杂交时要保证有足够的 DNA 浓度。DNA 在 50～100μl 杂交液/平方厘米滤膜为宜。

（4）杂交温度：高温下双链 DNA 变性成单链，温度降低，互补单链 DNA 又复性成双链，温度是影响杂交过程的一个重要因素。杂交温度一般控制在 Tm 值 - 25℃左右为宜，温度过高或过低，都会使分子的布朗运动速率减低，错配碱基间氢键不易断裂，杂交的特异性下降。变性温度计算经验公式

变性温度 = 81.5℃ - 2.76M +'0.41（G + C）% - 500/n - 0.61（甲酰胺 5%）

其中 M 为 Na^+ 的摩尔浓度；n 为探针的复杂性（没有重复序列时，即为探针长度（bp 数）。例如：一个长度为 500bp 的探针，（G + C）含量为 50%，杂交体系中含有 5 × SSC（0.75mol/L Na^+）和 50%甲酰胺，则其变性温度为：

81.5℃ - 2.07 + 20.5 - 1 - 30.5 = 68.4℃

那么，其杂交温度为 68.4℃ - 25℃ = 43.4℃。甲酰胺是 DNA 分子的变性剂，一般而言，有甲酰胺存在时杂交温度为 42℃，无甲酰胺存在时可用 65～68℃。杂交后最终洗膜温度一般应低于 Tm 值 5～12℃，一般在 55～62℃。

（5）杂交溶液的离子强度：在杂交混合物中，带有同性电荷的 DNA 分子互相排斥。同性电荷量越高，排斥越甚。在某一杂交缓冲液系统中，互补 DNA 链往往带有同种电荷，不利于杂交。为了消除电荷排斥，杂交液中应有较高浓度的盐，一般用 5 × 或 6 × SSC，增加溶液的离子强度，消除 DNA 的净电荷。

（6）杂交溶液的体积：溶液体积越大，杂交效率越低。在保证覆盖膜的前提下，应尽量减少杂交溶液的体积。一般每平方厘米膜加杂交液 50～100μl。

（7）杂交时间：一般为 $C_0t_{1/2}$的 1～3 倍；C_0 为当时间为 0 时的 DNA 浓度。$t_{1/2}$为核酸复性反应完成一半的时间。

（8）为减少非特异性杂交反应，在杂交前进行预杂交，将非特异性 DNA 结合点封闭，减少膜对探针的非特异性吸附作用。常用的封闭剂有两类：变性的非特异性 DNA，如鲑鱼精子或小牛胸腺 DNA 和高分子化合物如 Denhardt 氏溶液（含聚蔗糖 400、聚乙烯吡咯烷酮和 BSA）或脱脂奶粉、酪蛋白。

（9）硫酸葡聚糖（Dextran sulfate，MW500 000）能促进 DNA 链间的缔合。微粒表面可吸附 DNA 探针分子，从而使 DNA 接触面积增加，有利于杂交反应，在 10%硫酸葡聚糖中杂交速度可提高 10～100 倍。

（10）寡核苷酸探针混合物的杂交：在进行 SSOP 分析时，目前较为常用的是反向杂交，不同序列的寡核苷酸固定在尼龙膜上，与靶 DNA 进行杂交，不同序列的寡核苷酸的 G + C 含量不同，Tm 值有很大的差异，给杂交温度的选择带来困难。解决的办法是在杂交

体系中加入 3mol/L 氯化四甲胺（Me_4NCl，tetramethylammonium chloride）。Me_4NCl 可选择地与 A－T 对结合，并使其稳定性提高，Tm 值升高，使 A－T 对的解链温度趋同于 G－C 对，使寡核苷酸探针的 Tm 值只与探针的长度有关而与碱基序列无关。

根据不同要求，可采取不同的强度杂交条件要求。在温度较低时，探针不仅可与完全互补的特异性靶 DNA 序列相结合，同时也可与不完全互补 DNA 序列结合而形成杂交体。在低强度下，碱基对并不完全互补的杂交体亦可形成（图 6－10）。只有在高强度条件下，碱基完全互补的杂交体才稳定；影响强度的因素有甲酰胺浓度、温度和杂交液离子强度。低甲酰胺浓度、低温度和高离子浓度的条件下，杂交强度低，反之，则强度高。

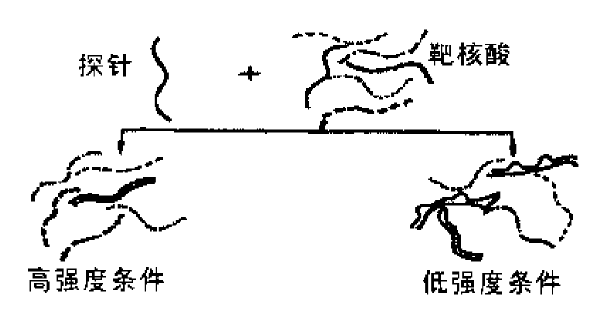


图 6－10 高、低强度条件下的杂交体

对于多基因座 DNA 指纹，既要检测出那些与探针完全匹配的靶 DNA 片段，也要检测出那些与探针碱基序列部分同源性的靶 DNA 片段，采用在高盐和低温下杂交以及在低洗涤强度下进行膜的漂洗。一般是在 0.5×SSC，0.1% SDS 缓冲液中 65℃或室温漂洗。单基因座 DNA 纹印采用高强度杂交条件，在低盐和高温下杂交以及在高洗涤强度下进行膜的漂洗。杂交操作在 DNA 指纹章节介绍。

需要注意的是，DNA 探针纯度影响到图谱的质量，DNA 探针中含有未被标记的 DNA 探针与标记的 DNA 探针竞争，使灵敏度降低，如果含有游离的 ^{32}P dNTP 或酶，将增高背景。

杂交、洗膜时间太短，溶液温度来不及变换，导致杂交强度条件不准确，在单基因座探针 DNA 图谱中会出现一些弱的非特异带。这种情况也会出现在盐浓度太高情况下，但温度过高或盐离子太低，又会使灵敏度降低。

在杂交、洗膜过程中，如摇动不够，会使杂交液和洗膜液与膜接触不均匀，自显影后出现信号强度不一致的情况。

膜表面易破损，尤其是在杂交过程中，增加背景，在所有过程中均应小心操作。另外，由于膜保存不善长真菌，损坏膜表面。因此要求膜用 1×SSC 冲洗后，彻底晾干后再置于室温或 4℃，如长期保存，需放于－20℃下。

尼龙膜韧性好，杂交过的膜通过洗去上次杂交的 DNA 探针，可以进行多次杂交，除去膜上结合探针的方法：将膜在 0.1% SDS 溶液中 80℃下洗 15min，再在 1×SSC 溶液漂洗一次，可以进行预杂交或晾干保存。

6.2.4 杂交信号的检测

6.2.4.1 放射性同位素标记的 DNA 探针系统

同位素标记 DNA 探针杂交后，利用同位素的放射线在 X 光片上成影作用检测杂交信号，称为放射自显影，所得到的 X 光片称为自显影（照）片。

6.2.4.2 非同位素标记 DNA 探针系统

除荧光和酶直接标记系统外，其他非放射性标记物系统不能直接被检测，需要经过偶联反应将酶与非放射性标记物偶联，再进行第二步的酶促反应来检测。根据酶促反应产物的颜色变化或化学反应发出的光来检测杂交结果。酶化学显色反应灵敏度低，目前仅用于斑点杂交中。常用的检测方法是化学发光技术。在氧化剂存在时，刚刚被氧化的发光底物常被辣根过氧化物酶或碱性磷酸酶催化氧化成激发状态，处于激发状态的分子极不稳定，要跃迁回基态。从激发态跃迁到基态时发出光，释放出的光能使 X 光片感光成影，显示 DNA 谱带。但光强度较弱。在反应体系中加入增强剂（酚类和荧光素类化合物，如羟基肉桂酸），使光子强度放大 400～1 000 倍，极大提高了仪器检测和胶片感光的灵敏度，其灵敏度接近同位素标记探针。

1. 辣根过氧化物酶

HRP 的显色底物有 DAB（3，3′－diaminobenzidine，二氨基联苯胺）和 TMB（3，3′，5，5′－tetramethylbenzidine，四甲基联苯胺）。它们在 HRP 作用下氧化脱氢直接在杂交部位产生一种不溶于水的有色物质。DAB 经 HRP 催化形成红色的产物，而 TMB 则为蓝色物。

辣根过氧化物酶以鲁米诺即氨基苯二酰肼（5－amino－2，3－dihydro－1，4－phthalazinedion）及其衍生物为发光剂，过氧化氢、过硼酸钠等过氧化物为氧化剂，催化反应发出光，在 X 光片上感光成像，记录 DNA 谱带（图 6－11）。用于反应中的增强剂（Enhancer）有 2 类：一类是 6－羟基苯噻唑及其衍生物，另一类是具有对位取代基的酚类，如对碘苯酚、对苯酚等。

对生物素标记的探针用链亲和素－HRP 作用，同理引起化学发光。

2. 碱性磷酸酶

碱性磷酸酶标记探针与化学发光底物 CSPD（disodium－3－(4－methoxyspiro {1，2－dioxetane－3，2′－(5′－chloro) tricyclo [3.3.1.1.3.7decan} －4－yl) phenyl phosphate）作用时，碱性磷酸酶使 CSPD 脱磷形成不稳定的中间产物，迅速分解并发光（477nm），感光 X－光片。3－（2′－螺旋金刚烷）－4－甲氧基－4－（3′－磷酰基氧化）酚－1，2－氧环化物，简称 AMPPD 及 Lumi－Phos530，是碱性磷酸酶的另一种发光底物，其结构及原理见图 6－12。

3. 偶联反应

半抗原的标记物如地高辛标记物，通过抗原－抗体免疫反应系统与显色或发光反应偶联。

生物素标记物是非放射性标记系统较为常用的标记系统。生物素可以作为抗生物素蛋白（卵白素，avidin）的配体，通过亲合素与酶偶联进行检测。Avidin 是一种糖蛋白，有 4 个与生物素结合的基因座，与生物素的亲合力极高。但是，由于 avidin 为糖蛋白，等电点偏碱，中性环境下带正电荷，易与其他带副电荷的生物大分子非特异性结合，再加上体内

Luminol $+ H_2O_2 \xrightarrow{辣根过氧化酶} + 2H^+ + 2H_2O + N_2 +$ 光子

标记探针-复合物以共价键与DNA结合
杂交
探针DNA结合到靶DNA上
检测试剂1　$O_2^- + H_2O$
检测试剂2 鲁米诺　氧化产物
增强剂
发光
X光片结果

图 6-11　辣根过氧化物酶催化化学发光反应

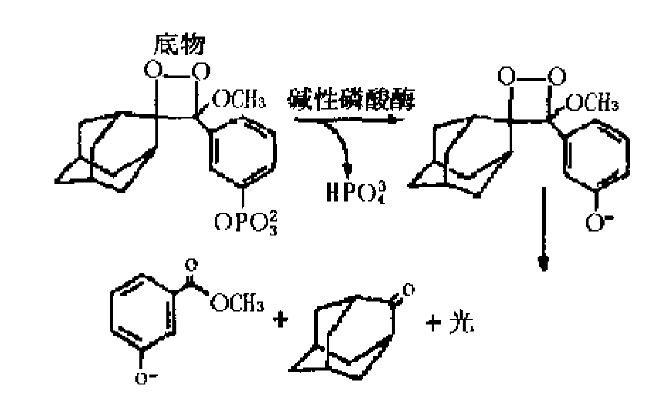


图 6-12　辣根过氧化物酶、碱性磷酸酶催化化学发光反应

存在的内源性生物素的干扰，易导致假阳性结果。因此目前大多数使用链亲和素（streptavidin），它不是糖蛋白，等电点为中性，较少形成非特异性结合，特异性极高。

6.2.5 原位杂交

原位杂交（Hybridization in situ）由 Granstein 和 Hogness 于 1975 年提出，其制样方法与 Southern blot 完全不同，利用 DNA 探针与组织切片或细胞中的相应 DNA 进行复性杂交而加以检测，称为原位杂交技术。这是一种分子杂交与组织化学相结合的技术，也称杂交组织化学、细胞杂交或原位组织化学。有以下特点：

1. 特异性高，可以精确定位。

2. 能在成分复杂的组织中进行单一细胞的杂交，而不受同一组织其他成分的影响。

3. 不需要从组织中提取 DNA，可完整地保持组织与细胞的形态，对于组织中含量极低的靶序列有极高的敏感性。

根据 DNA 探针标记物是否能直接被检测，原位杂交可分为直接法和间接法两类，图 6-13 为直接法与间接法原位杂交示意图。原位杂交技术广泛应用于医学生物学，其基本程序包括：组织细胞标本的制备→预杂交→杂交→漂洗→放射自显影或免疫酶联法显色。在法医物证检验中应用较少，曾用于检测 Y 染色体 DNA 进行性别检验。目前主要在病理学上用特异性病毒核酸分子作为探针进行原位杂交。检测组织细胞中病毒 DNA 或 RNA，以确定病毒感染，或检测病理损伤的 DNA。

图 6-13　位杂交示意图

a：直接法 b：间接法

第七章 聚合酶链式反应

聚合酶链式反应（polymerase chain reaction，PCR），是近年发展起来的一种体外扩增特异 DNA 片段的技术。体外扩增的思想早在 1971 年就由 Khorana 及同事提出，在 1985 年由 Kary Mullis 创立。此法操作简便，短时间内在试管中可获得数百万个特异 DNA 序列的拷贝。1993 年 Mullis 因此获得了诺贝尔奖。PCR 技术已迅速渗透到分子生物学的各个领域，在分子克隆、遗传病的基因诊断、法医学、考古学等方面得到了广泛的应用，而且与 PCR 相关的技术不断被开发应用。目前大多数法医 DNA 分析技术均以 PCR 技术为基础进行分型。

7.1 PCR 技术的原理

PCR 技术实际上是模拟生物体内的在模板 DNA、引物和 4 种脱氧核糖核苷酸存在下，依赖于 DNA 聚合酶的体外 DNA 酶促合成反应。PCR 技术的特异性取决于引物和模板 DNA 结合的特异性，引物是 PCR 反应中最主要的成分。反应分三步：（1）变性（denaturation）：首先是双链 DNA 分子在 95℃左右高温下加热，DNA 双螺旋配对碱基间的氢键断裂，双链解离成单链 DNA；（2）退火（annealing）：降低温度使引物和其互补的模板形成杂交链。由于模板分子结构较引物要复杂的多，而且反应体系中引物 DNA 量大大多于模板 DNA，模板 DNA 双链之间互补的机会较少；（3）延伸（extension）：在 DNA 聚合酶和 4 种脱氧核糖核苷酸底物及 Mg^{2+} 存在的条件下，聚合酶催化以引物为起始点的 5′→3′方向的 DNA 链延伸反应。

经过高温变性，低温退火和中温延伸 3 个温度的循环，模板上介于两个引物之间的序列不断得到扩增。每循环一次，目的 DNA 的拷贝数加倍，随着循环次数的增加，目的 DNA 以指数的形式堆积，数量可达 $2 \times 10^{5\sim7}$ 拷贝。在反应的最初阶段，原来的 DNA 担负着起始模板的作用，随着循环次数的递增（图 7－1），由引物引导延伸的片段急剧地增多而成为主要模板。因此，绝大多数扩增产物将受到所加引物 5′末端的限制，最终扩增产物是介于两种引物 5′端之间的 DNA 片段。

7.2 PCR 反应

标准的 PCR 反应一般选用 50～100μl 反应体系，其中含有：

50mmol/L KCl

10mmol/L Tris－HCl（室温下 pH 8.4）

1.5mmol/L $MgCl_2$

0.001%明胶或牛血清白蛋白（BSA）

2 个引物，各 0.25μmol/L

4 种底物（dATP＋ dCTP＋ dGTP＋ dTTP），各 200μmol/L

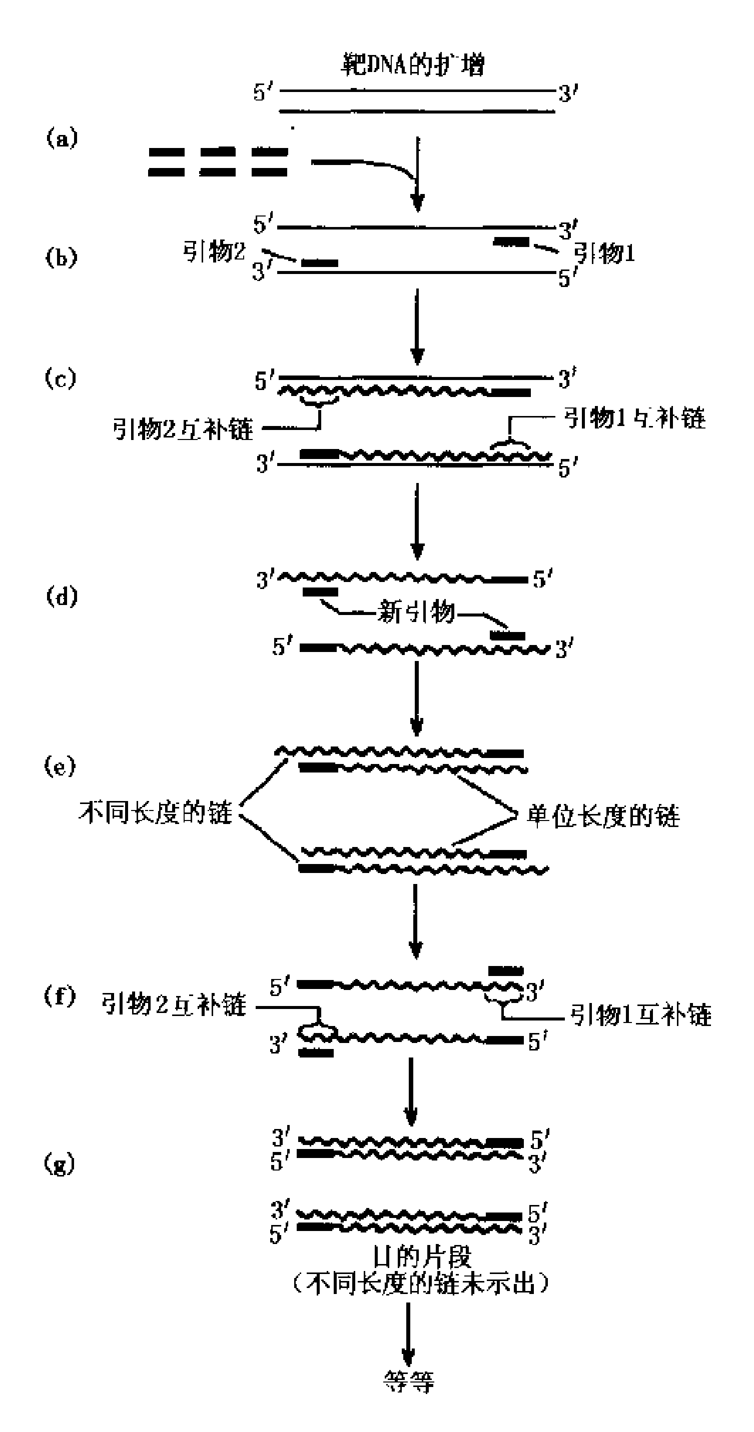


图 7-1　聚合酶链式反应示意图

(a) 作为起始材料的双链 DNA 分子；(b) 反应混合物加热后发生链的分离，然后降温使引物结合到位于待扩增的靶 DNA 区域两端的退火位置上；(c) Taq 聚合酶以单链 DNA 为模板在引物的引导下利用反应混合物中的 dNTPs 合成互补的新链 DNA；(d) 将反应混合物再次加热，使旧链和新链分离开来；这样便有 4 个退火位置可供引物结合，其中两个在旧链上，两个在新链上；(e) Taq DNA 聚合酶合成新的互补链 DNA，但这些链的延伸是精确地限定于靶 DNA 序列区。因此，这两条新合成的 DNA 链的长度是严格地定位在两条引物界定的区段内；(f) 重复过程，引物结合到新合成的 DNA 单链的退火位点；(g) Taq 聚合酶合成互补链，产生两条与靶 DNA 区段完全相同的双链 DNA 片段。

模板 DNA（人基因组 DNA）0.1μg

Taq DNA 聚合酶 2.5u

反应条件一般为 94℃变性 30s，55℃退火 30s，70～72℃延伸 30～60s，30 个循环。

开始进行 PCR 试验时，一般以标准 PCR 反应为基础，逐步改变和优化热循环和反应体系组成条件，直至建立一个最适的 PCR 反应。

7.3 PCR 反应的优化

PCR 必须具备下述基本条件：1. 模板 DNA；2. 人工合成的寡核苷酸引物；3. 合适的缓冲体系；4. Mg^{2+}；5. 三磷酸脱氧核苷酸；6. 耐热 DNA 聚合酶；7. 温度循环参数（变性、复性和延伸的温度与时间以及循环数）。另外，还有一些其它因素，如二甲基亚砜、甘油、石蜡油、明胶或小牛血清白蛋白等对某些特定 PCR 有一定的影响作用。

7.3.1 模板 DNA

PCR 是以 DNA 为模板进行的核酸体外扩增。法医 DNA 分析的 DNA 标本来源广泛，种类多，如血斑、毛发、精斑、新鲜的或经甲醛固定或石蜡包埋的组织以及陈旧骨骼等，条件不一。PCR 技术对模板 DNA 的要求是不能混有蛋白酶、核酸酶、Taq 酶抑制剂以及能与 DNA 发生结合的蛋白质等，因为这些物质会影响 PCR 扩增，另外，残留过高浓度的盐或 EDTA 都可能抑制 PCR 反应。关于 DNA 的具体处理方法详见 DNA 提取部分。

PCR 反应中模板加人量一般为 10^2～10^5 拷贝的靶序列。1μg 人基因组 DNA 相当于 3×10^5 个单拷贝的靶分子。理论上，反应中仅需加人 0.5～2ng 人基因组 DNA 即可。

扩增靶序列的长度根据不同目的而不同。虽然用 Taq 聚合酶在合适条件下和较长延伸时间下，可扩增长达 10～20kb 的片段，但一般在 500bp 以下，以 100～300bp 为最好。

7.3.2 引物

PCR 扩增产物的长度、特异性是由特异引物限定的。因此，引物的设计与合成是保证 PCR 成功的关键因素之一。

1. 引物合成的质量

合成的引物中会含有一些“错误序列”，其中包括不完整的序列和脱嘌呤产物等，这些序列可导致非特异扩增和降低目的片段扩增效率。因此，合成的引物必须经聚丙烯酰胺凝胶电泳或反向高压液相层析（HLPC）纯化，以达到只含有单一的引物序列。反复冻融可引起引物的降解，引物应少量分装保存，避免反复冻融。冻干引物在 -20℃可保存 12～24个月，液体状态在 -20℃可保存 6 个月。

2. 引物的设计原则

引物的优劣直接关系到 PCR 的特异性与成功与否，因此，引物的设计应遵循以下原则。

（1）长度

寡核苷酸引物长度为 15～30bp，一般为 20～27mer。引物的有效长度

（Ln）=2×（G+C）+（A+T），Ln 值不能大于 38，如果 Ln>38，最适延伸温度会超过 Taq DNA 聚合酶的最适温度（74℃），不能保证产物的特异性。

（2）G十C含量

G+C 含量一般为 40%～60%。寡核苷酸的解链温度（Tm 值）与 GC 含量有关。

$Tm=4(G+C)+2(A+T)$，有效引物的 Tm 为 55～80℃。

（3）碱基的随机分布

引物中四种碱基的分布应当是随机的，不要有聚嘌呤或聚嘧啶的存在。尤其3′端不应超过3个连续的G或C，因这样会使引物在G+C富集序列区引发错误。

（4）引物自身碱基结构

引物自身不应存在互补序列，否则引物自身会折叠成发夹状结构或引物本身复性。这种二级结构会因空间位阻而影响引物与模板结合。引物自身连续互补碱基不能超过3对。另外，两引物在3′端不能出现同源序列，否则会形成引物二聚体；两引物之间不应有互补性，一对引物间不应有多于4个连续的同源性或互补性碱基。

（5）引物的3′末端碱基

引物的延伸是从3′端开始的，引物3′端的末位碱基在很大程度上影响 Taq DNA 聚合酶的延伸效率，不能进行任何修饰。3′端也不能有形成任何二级结构的可能，除在特殊的 PCR（AS－PCR）反应中，引物3′端不能发生错配。

有资料表明，引物3′端不同碱基扩增引发效率有很大差异（表7－1），3′端碱基应选用T、C、G，以T为最佳，以3′端AT含量多，且5′端及中端GC含量多为好，可减少非特异扩增。

表7－1　引物3′端碱基错配对扩增效率的影响

		引物3′端碱基			
		T	C	G	A
模板3′端碱基	T	1.0	1.0	1.0	1.0
	C	1.0	≤0.01	1.0	1.0
	G	1.0	1.0	1.0	≤0.01
	A	1.0	1.0	≤0.01	0.05

注：完全配对产物量定为1.0

同等条件下，引物3′端第2位或第3或4位的错配对扩增量影响不大，但易出现非特异扩增和引物二聚体。

（6）引物的5′端碱基

引物的5′端限定着 PCR 产物的长度，它对扩增特异性影响不大。因此，可以被修饰而不影响扩增的特异性。引物5′端修饰包括：加酶切位点；标记生物素、荧光、地高辛、Eu^{3+}等；引入蛋白质结合 DNA 序列；引入突变位点、非人基因组 DNA 序列（如 MVR 分析中 TAG 序列）等。

（7）密码子的简并

在扩增编码区域时，引物3′端不要终止于密码子的第3位碱基，因为密码子的第3位易发生简并，会影响扩增特异性与效率。

（8）引物的特异性

引物与非特异扩增序列的同源性不要超过70%或有连续8个互补碱基同源。

(9) 避开引物的二级结构区

某些引物无效的主要原因是引物重复区 DNA 二级结构的影响，选择扩增片段时最好避开二级结构区域。若不能避开这一区域时，用 7 - deaza - 2′ - 脱氧 GTP 取代 dGTP 对扩增的成功是有帮助的。

3. 引物的用量及其计算

一般 PCR 反应中引物的终浓度为 0.1 ~ 1μmol/L，在此范围内，PCR 产物量基本相同。引物低于 0.1μmol/L 时，则产量降低。引物浓度过高会促进引物的错误引导合成非特异产物，增加引物二聚体的形成。如果错配或引物二聚体出现在早期的循环中，也成为 PCR 反应的模板，与靶序列竞争 DNA 聚合酶、dNTP 底物，从而使靶序列的扩增量降低。不同的反应引物的最适浓度也不同，最好选用几种不同浓度实验再作调整，尽可能降低引物浓度，对整个 PCR 反应是有益的。

摩尔消光系数（E_M）是 1cm 光程比色杯中测定 1mol/L 寡核苷酸溶液在 UV_{260nm} 下的光密度（OD）值。E_M 可按下式计算：

$E_M = a（16\,000）+ b（12\,000）+ c（7\,000）+ d（9\,600）$

其中，a、b、c 和 d 分别代表寡核苷酸中 A、G、C 和 T 的个数。

例如：一纯化的 20mer 寡核苷酸溶于 0.1ml 水中，取 10μl 稀释至 1.0ml，测其 OD = 0.50，则

原液的 OD 为 $0.50 \times 100 = 50$，若此寡核苷酸碱基组成为 A = 5，G = 5，C = 5，T = 5，其 EM 为：

$E_M = 5（16\,000）+ 5（12\,000）+ 5（7\,000）+ 5（9\,600）= 223\,000$

原液中寡核苷酸的摩尔浓度为：

$50 \div 223\,000 = 2.9 \times 10^{-4}mol/L = 290nmol/L$

也可以用小卫星 - PCR 章节介绍的方法计算。

7.3.3 缓冲液

标准缓冲体系为 10 ~ 50mmol/L Tris - HCl（pH8.3 ~ 8.8，-20℃）。Tris 是一种双极性离子缓冲液，改变反应液的缓冲能力，会引起 PCR 扩增效率的改变。如将 Tris 浓度增加到 50mmol/L（pH8.9），有时会增加产量。

反应混合液中 KCl 浓度在 50mmol/L 下有利于引物的退火，超过 50mmol/L 的 KCl 则抑制 Taq DNA 聚合酶的活性。有些反应液中以 NH_4^+ 代替 K^+，其浓度为 16.6mmol/L。反应中加入小牛血清白蛋白（100μg/ml）或明胶（0.001%）或 Tween - 20（0.05% ~ 0.1%）有助于酶的稳定，反应中加入 5mmol/L 的二硫苏糖醇（DTT）也有类似作用，尤其在扩增长片段或有抑制剂存在时，加入这些酶保护剂对 PCR 反应是有利的。

7.3.4 Mg^{2+}

Mg^{2+} 是 Taq DNA 聚合酶活性所必需的。因此，反应中优化 Mg^{2+} 浓度显得十分重要。Mg^{2+} 浓度除影响酶活性与忠实性外，也影响着引物的退火、模板与 PCR 产物的解链温度、产物的特异性和引物二聚体的形成等。Mg^{2+} 浓度过低时，酶活力显著降低；过高时，则酶催化非特异的扩增。PCR 反应中 Mg^{2+} 浓度在 0.5 ~ 2.5mmol/L 之间。需注意的是，Taq DNA 聚合酶需要的是游离的 Mg^{2+}，而 PCR 混合物中的 DNA 模板、引物和 dNTP 的磷酸基

团均可与 Mg^{2+} 结合，降低 Mg^{2+} 实际浓度。因此，PCR 中 Mg^{2+} 的加入量要比 dNTP 浓度高 0.2~2.5mmol/L。另外还须注意引物和模板 DNA 中如含 EDTA 等螯合剂也会影响游离 Mg^{2+} 浓度。

最佳 Mg^{2+} 浓度确定方法：在模板 DNA 量，引物和 dNTP 浓度和设定的 PCR 循环参数确定情况下，反应中使用不加 Mg^{2+} 的 PCR 缓冲液中。Mg^{2+} 浓度按 0.5mmol/L 递增（0.5，1.0，1.5，2.0，2.5~5.0mmol/L）实验。在确定了 Mg^{2+} 大概浓度以后，再在该浓度上下，以 0.2mmol/L 递增和递减数个浓度来确定 Mg^{2+} 的最适浓度。

7.3.5 三磷酸脱氧核苷酸（dNTP）

反应中每种 dNTP（dATP，dCTP，dTTP 和 dCTP）的终浓度为 20~200μmol/L，在此范围内，PCR 产物量、特异性与合成忠实性间的平衡最佳。所用的四种 dNTP 终浓度相等，可使错误掺入率降至最低。

4 种 dNTP 的浓度对保持碱基掺入忠实性是很重要的。dNTP 终浓度大于 50mmol/L 时会抑制 Taq DNA 聚合酶活性。最低的适宜 dNTP 浓度可根据特定靶序列长度和碱基组成来确定。如在 100μl 反应液中，每种 dNTP 若为 20pmol/L，理论上足以合成 2.6μg 产物。

贮备 dNTP 液应用 NaOH 调 pH 至中性，浓度用分光光度计测定。贮备液为 5~10mmol/L，分装后 -20℃保存。

另外，dNTP 的类似物也可掺入到 PCR 产物。

7.3.6 耐热 DNA 聚合酶

在 PCR 反应中，DNA 聚合酶是最关键的因素之一。PCR 发明的初期使用的 DNA 聚合酶是大肠杆菌 DNA 聚合酶 I 的 Klenow 片段。由于：（1）该酶的热稳定性差，每次 DNA 热变性后，该酶绝大部分已失去活性，每一至数个循环即需添加新酶；（2）Klenow 酶反应温度偏低，容易形成引物与 DNA 模板的非专一性配对，产生非特异性扩增。由于温度低，反应易受某些 DNA 二级结构的影响而得不到完整的扩增产物。Klenow 酶的热稳定性差和使用不便限制了其在 PCR 反应中的应用。以后人们曾使用 T4 DNA 聚合酶和 T7 DNA 聚合酶，两者虽使扩增特异性增加和 DNA 合成速度提高，但终因对热的稳定性差而未得到广泛应用。直到耐热 DNA 聚合酶（如 Taq DNA polymerase）应用于 PCR 反应，才使这一技术得到迅速发展和广泛的应用。

Taq DNA 聚合酶是从一种嗜热水生菌 Thermus aquaticus YT-1 菌株中分离提纯的。它是从美国黄石国家森林公园火山温泉中分离得到的，生长在 70~75℃极富矿物质的环境中。在嗜热脂肪芽孢杆菌及其他一些嗜热菌中也得到类似的耐热 DNA 聚合酶。

Taq DNA 聚合酶基因全长 2496 个碱基，编码 832 个氨基酸，其酶蛋白分子量为 94kD。该酶在 70~75℃时具有最高的生物学活性（图 7-2a）。75~80℃条件下每一个酶蛋白分子每秒可延伸约 150 个核苷酸。温度降低时，Taq DNA 聚合酶延伸速度明显下降，如 70℃时，酶分子延伸速率在 60 个核苷酸/秒以上，55℃时为 22 核苷酸/秒；而 37℃和 22℃时分别为 15 核苷酸/秒和 0.25 核苷酸/秒。当温度超过 80℃时，合成速度也明显下降，可能是引物或引物-模板的稳定性遭到了破坏。Taq DNA 聚合酶最适 pH 值为 8.3~9.0（图 7-2b）。

Taq DNA 聚合酶具有良好的热稳定性。实验结果表明 92.5℃、95℃、97℃下 DNA 聚合

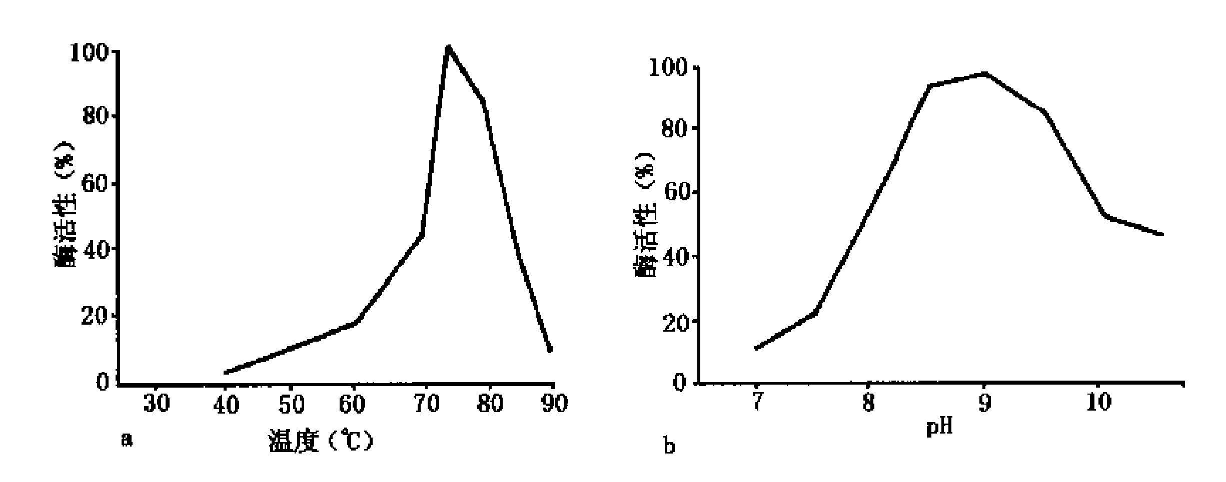


图 7-2 温度、pH 值对 Taq DNA 聚合酶活性的影响

a：温度对 Taq DNA 聚合酶活性的影响；b：pH 值对 Taq DNA 聚合酶活性的影响

酶的生物半衰期分别为 130min、40min 和 5～6min。在 PCR 反应中，变性温度一般为 94℃ 30s，最长不超过 60s，以循环 30 次计算，Taq DNA 聚合酶可保证 PCR 反应的需要（图 7-3）。

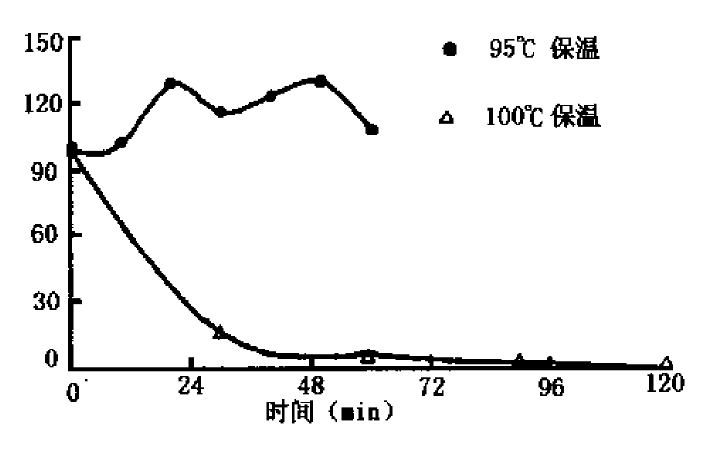


图 7-3 Taq DNA 聚合酶热稳定性

与大肠杆菌 DNA 聚合酶Ⅰ相比，Taq DNA 聚合酶具有 5′→3′聚合酶活性和 5′→3′外切酶活性，没有 3′→5′外切酶活性。因此，在 PCR 反应中如果发生某些单核苷酸的错配，这种酶是没有校正功能的。使用 Taq DNA 聚合酶的 PCR 反应出现碱基错配的机率为 2.1×10^{-4}左右。

和许多其它聚合酶一样，Taq DNA 聚合酶是 Mg^{2+} 依赖性酶，因此，Taq DNA 聚合酶活性对 Mg^{2+} 浓度非常敏感。以鲑鱼精 DNA 为模板，dNTP 的总浓度在 0.7～0.8mmol/L，在不同浓度 $MgCl_2$ 进行 PCR 反应 10min，结果表明 $MgCl_2$ 浓度在 2.0mmol/L 时酶活力最高，10mmol/L $MgCl_2$ 使酶活性下降 50%。Mg^{2+} 浓度偏高，酶活性受抑制。由于 Mg^{2+} 能与负离子或负离子团（如磷酸根等）结合，在 PCR 反应中模板 DNA、引物及 dNTP 均可与 Mg^{2+} 结合，尤以 dNTP 的影响最大。所以在不同反应体系中 $MgCl_2$ 浓度应适当进行调整。一般反应中 Mg^{2+} 浓度至少应比 dNTP 总浓度高出 0.2～2.5mmol/L。

KCl最适浓度是50mmol/L，高于75mmol/L时，PCR反应明显抑制。均衡的dNTP有利于酶活性发挥，并能减少错配和获得更多的特异性DNA扩增产物。不同变性剂对酶活性的影响不一样（表7-2），了解这些特性有助PCR反应的设计、优化。

表7-2　变性剂对TaqDNA聚合酶活性的影响

抑制剂	浓度	活性（%）
乙醇	≤3%	100
	10%	10
尿素	≤0.5mol/L	103
	1.0mol/L	118
	1.5mol/L	107
	2.0mol/L	82
DMSO	≤1%	100
	10%	53
	20%	11
DMF	≤5%	100
	10%	82
	20%	17
甲酰胺	≤10%	100
	15%	86
	20%	39
SDS	0.001%	105
	0.01%	10
	0.1%	<0.1

自从耐热Taq DNA聚合酶引入PCR后，又发现了多种耐热的DNA聚合酶，有多种耐热DNA聚合酶（VENT和Tth等）相继用于PCR。其中，Vent^{TM} DNA聚合酶是美国New England Biolabs公司从海底火山口中在近100℃高温下生长的Thermoccus litoralis菌中提纯得到的，性能比Taq DNA聚合酶更优越的另一个耐热酶。该酶在100℃时半衰期达95min，此外还具有3′→5′外切酶活性。但目前仍以Taq DNA聚合酶应用较多。不同来源聚合酶制备条件、测定方法、单位定义有所不同，使用时需加以注意。

Gold Taq DNA聚合酶是Taq DNA经过化学修饰的酶，热稳定性、特异性好，酶延伸反应需在高温（95℃）激发。

7.3.7 温度循环参数

1. 变性温度与时间

PCR的变性这一步很重要。此步若不能使靶基因模板和（或）PCR产物完全变性，就会导致PCR失败。DNA在其链分离温度（strand separation temperature，Tss）时变性只需几秒钟，但反应管内部达到Tss还需一定时间。典型的变性条件是95℃ 30s，或97℃15s，更高的温度可能更有效，尤其是对富含G+C的靶基因。但变性温度太高会影响酶活性。最简单有效的方法是在加Taq聚合酶前先使模板在95℃变性7~10min，在以后的循环中，将模板DNA在94℃或95℃变性1min。在扩增短片段（100~300bp）时，可用简便、快速

的两步 PCR 法，同时在 5～10 个循环后，将变性温度降至 87～90℃可改善 PCR 产量，具体降低程度则根据具体反应和使用的 PCR 仪而定。

2. 退火温度与时间

变性温度是 PCR 反应的成败关键，退火温度决定着 PCR 的特异性。变性后 DNA 很快冷却，可使引物和 DNA 结合。引物退火温度与时间取决于引物的碱基组成（G + C 含量）、长度和浓度。长度在 15～25bp，引物退火温度 Tm = 4（G + C）+2（A + T），引物过长，计算所得 Tm 会偏高。合适的退火温度应比 Tm 低 5℃。实际反应中可能应以较 Tm 低 6～8℃开始为宜，逐渐增加，每次 2℃。

选择合适的退火温度可以增高反应的产率，大大减少引物与模板的非特异结合，提高 PCR 反应的特异性。由于反应体系中，初始状态下存在过量的引物，在开始几个循环选用稍高一点的温度，可能有助于提高反应的特异性。表 7－3 为理论计算的碱基组成与 PCR 引物退火温度关系，从表中可根据 AT 与 GC 的个数，从其交叉点迅速查出较为合适的退火温度。有资料表明根据此表指导分析 50 对引物 PCR 的最适退火温度，其中有 49 对引物的实验值与计算值相差在 4℃以内。

当引物长度为 12～15bp 时，Tm 值可能只有 40～45℃。这样的引物不可能在 72℃保持退火状态。因此，这类反应需要利用 Taq 酶在较低温度下的活性，将反应在 50～60℃保温或将温度缓慢地升到 72℃来完成整个反应。

表 7－3 寡脱氧核苷酸的的 Tm

(GC)→ (AT)↓	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
3									61	64	66	69	71	73	75	76	78	79	81	82	84	85	86
4								59	61	64	66	69	71	73	75	76	78	79	81	82	84	85	86
5							56	59	61	64	66	69	71	73	75	76	78	80	81	82	84	85	86
6						53	56	59	62	64	67	69	71	73	75	76	78	80	81	83	84	85	
7					50	53	56	60	62	64	67	69	71	73	75	77	78	80	81	83	84		
8				47	51	54	57	60	62	65	67	69	71	73	75	77	78	80	81	83			
9			44	48	52	55	57	60	63	65	68	70	72	74	75	77	78	80	81				
10		41	45	49	53	55	58	61	63	66	68	70	72	74	75	77	78	80					
11	39	43	46	50	53	56	59	61	64	66	68	70	72	74	75	77	80						
12	40	44	48	51	54	57	60	62	64	66	69	71	72	74	75	77							
13	42	46	49	52	55	58	60	63	65	67	69	71	72	74	75								
14	43	47	50	53	56	59	61	63	65	67	69	71	73	74									
15	44	48	51	54	57	59	61	64	66	68	70	71	73										
16	45	49	52	55	57	60	62	64	66	68	70	72											
17	47	50	53	56	58	61	63	65	67	69	70												
18	48	51	54	57	59	61	63	65	67	69													
19	49	52	55	58	60	62	64	66	68														
20	50	53	56	58	60	62	64	66															
21	51	54	57	59	61	63	65																
22	52	55	57	60	62	64																	
53	56	58	60	62																			
24	54	57	59	60																			
25	55	58	60																				

确定了退火温度后，退火时间并不是关键因素。但退火时间太长会增加非特异的复性，太短了产量低，一般不宜少于30s，以保证引物与模板DNA结合。

3. 延伸温度与时间

引物延伸温度一般为72℃（较复性温度高10℃左右），此时Taq DNA聚合酶活性最高。核苷酸的掺入率为3.5~100个核苷酸/s，具体取决于缓冲体系、pH、盐浓度和DNA模板的性质。不合适的延伸温度不仅会影响扩增产物的特异性，也会影响其产量。

延伸时间决定于靶序列的长度与浓度，对于长达2kb的扩增片段，72℃延长1min足够了，3~4kb的靶序列则需3~4min延伸。PCR中前几个循环延伸时间应足够长，使靶序列延伸完全。对浓度很低模板的扩增，延伸时间要长些。但延伸时间过长会导致非特异扩增带的出现。

4. 循环参数

循环数决定着扩增程度。过多的循环会增加非特异扩增产物，过少的循环数，PCR产物量就会极低。在其它参数都已优化条件下，最适循环数取决于靶序列的初始浓度。表7-4列出了循环数与靶DNA片段PCR扩增关系。在初始靶序列为 3×10^5，1.5×10^4，1×10^3 和50拷贝分子时，循环数可分别为25~30，30~35，35~40和40~45个循环。

表7-4 循环数与扩增产物

循环数	靶DNA分子数	循环数	靶DNA分子数
1	0	17	32768
2	0	18	65536
3	2	19	131072
4	4	20	262144
5	8	21	524288
6	16	22	1048576
7	32	23	2097152
8	64	24	4194304
9	128	25	8388608
10	256	26	16777215
11	512	27	33544432
12	1024	28	67108864
13	2048	29	134217728
14	4096	30	268435456
15	8192	31	536870912
16	16384	32	1073741824

7.3.8 PCR促进剂

1. 二甲基亚砜（DMSO）：许多耐热DNA聚合酶厂家推荐在PCR反应中加10%DMSO，这可能是DMSO有变性DNA的作用。DMSO的使用对大肠杆菌DNA聚合酶Ⅰ Klenow片段是有益的，但对Taq DNA聚合酶有抑制作用，一般反应中应尽量不用DMSO。在复合PCR中有时选用。

2. 甘油：有报道，反应中加5%~20%甘油有助于PCR反应的复性过程，尤其对GC

含量高和二级结构多的靶序列以及扩增长片段（>1500bp）更适用。应注意的是，DMSO和甘油并非对所有 PCR 均有益，因此，是否加这些试剂应根据具体情况而定，也需要操作者的摸索。

3. 氯化四甲基铵（TMAC）：在反应中加入 $1\times10^{-5}\sim1\times10^{-4}$mol/L 的 TMAC 可促进 PCR，去除非特异扩增，而不抑制 Taq DNA 聚合酶。

4. T4 噬菌体基因 32 蛋白质（gp32）：加入 0.5~1μl gp32 可使 Taq DNA 聚合酶对长片段 DNA 的扩增改善至少 10 倍。

5. 石蜡油：反应中所用石蜡油质量要高，不能含抑制 Taq DNA 聚合酶活性的杂质。加石蜡油目的是防止反应被蒸发后引起反应成分的改变。石蜡油的有无对反应影响较大。但对于有热盖的 PCR 仪，如 PE—Cetus 公司的 2400、9600、9700 等型基因扩增 PCR 系统，可以不加石蜡油。

7.3.9 PCR 仪

自第一台 PCR 仪问世以来，已有许多热循环 PCR 仪相继诞生，不同仪器由于温控精度不同对 PCR 反应有一定影响。因此，优化 PCR 反应时要注意不同仪器间的差别。

7.3.10 平台效应

经过一定次数的热循环，PCR 产物渐渐停止呈指数状态的增加，呈线性增加或静止状态的这个时期被称为“平台期”，进入平台期的这种现象叫平台效应（plateau effect），此时有0.3~1pmol靶序列的累积。PCR 反应达到平台期的时间，主要取决于反应开始前靶 DNA 的拷贝数与 DNA 合成总量。根据反应条件和热循环参数，下列因素可能与平台效应有关：(1) dNTP 和引物快速掺入底物中，浓度降低；(2) 随产物的增加酶与模板的比例下降；(3) 由于变性温度高（93~95℃）和温度循环，酶活力和 dNTP 的稳定性逐渐下降；(4) 非特异产物或引物二聚体与反应物的竞争；(5) 产物在高浓度时变性不完全，影响引物的延伸；当浓度高于 10^{-8}mol/L 时，可能降低 Taq 聚合酶的延伸，甚至导致产物特异性下降，酶与 PCR 产物的结合，使酶分子减少。

7.3.11 忠实性

PCR 扩增过程能否按照模板忠实地扩增是个重要问题，诱发错误核苷酸掺入的条件是 dNTP 浓度明显低于 1μmol 或四种 dNTP 浓度不平衡，即其中一种 dNTP 浓度相对低于其他三种。由于 Taq DNA 聚合酶无 3′→5′外切酶活性，错误掺入的碱基不能被校正，但错配的碱基不能有效地延伸，限制了无效分子的扩增，有利于保持扩增的忠实性。采用高于 55℃的退火与延伸温度，配合适当的平衡 dNTP 浓度（每种 dNTP 10~50μmol/L）可提高 PCR 终产物的忠实性。

在扩增过程中存在极少量的核苷酸掺入错误。PCR 产物中一部分含有错误掺入，根据计算和实验证明，不会影响遗传分型。Taq DNA 聚合酶掺入错误速度为 $1\sim2/10^4$ 掺入。按这个速度，在 DNA 序列特定位置发生特定的错误掺入机会非常少，不会发生基因型改变，即使发生，也很难检测到，除非特定错误掺入的量为 PCR 产物的 5%。

PCR 产物中特定错误掺入比例受二个因素的影响：模板 DNA 拷贝数和 PCR 循环数（表 7-5）。如果起始模板只有几个拷贝，在第一个循环发生错误掺入，就会发生最糟的事，即 25%PCR 产物是错误掺入的。当扩增模板很多或错误掺入发生在后面的循环中，

特定错误掺入比例随之减少。

表 7－5　模板与 PCR 循环对核苷酸错误掺入的影响

起始模板 DNA 拷贝数	错误取代循环			
	1	2	5	r
1	0.25	0.125	0.015	$1/2^{r}+1$
10	0.025	0.0125	0.0015	$0.1/2^{r}+1$
100	0.0025	0.00125	0.00015	$0.01/2^{r}+1$

7.4 PCR 相关技术

PCR 技术发展速度，不断产生一些新的 PCR 相关技术，下面就这些技术作简单介绍。

7.4.1 PCR－SSO 技术

1. 原理

序列特异性寡核苷酸－聚合酶链反应（sequence specific oligonucleotide－PCR，SSO－PCR）或等位基因特异寡核苷酸－聚合酶链反应（allelic specific oligonucleotide PCR，PCR－ASO）是利用合成与一种或多种等位基因特异碱基序列的互补寡核苷酸探针，与固定在尼龙膜上的 PCR 扩增产物进行杂交，检测有碱基序列差异的特定等位基因多态性（图 7－4）。PCR－SSO 方法早在 1986 年由 Saik 等人报道，采用该技术检测 HLA－DQα 多态性和 β－球蛋白基因。同年，Amgelini 等人报道了用 PCR－SSO 技术检测 HLA－DR 基因多态性，1988 年 Bugawan 等人利用 PCR－SSO 技术分析 HLA 基因多态性并将之应用于法医学检验中。

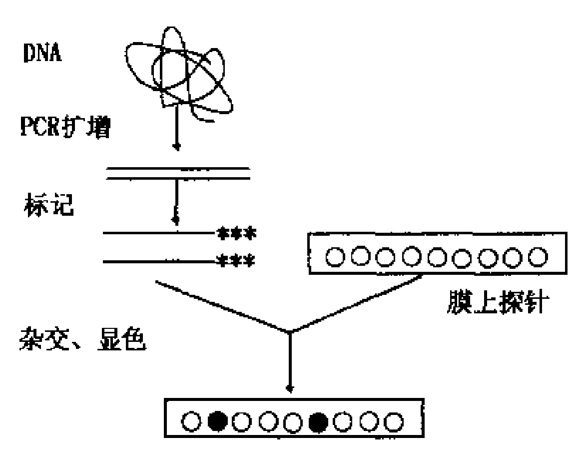


图 7－4　PCR－SSO 分析原理示意图

PCR－SSO 技术主要包括提取目的基因 DNA，以位点间或组间特异性引物进行体外扩增，再将扩增产物转移到硝酸纤维素膜或尼龙膜膜上，然后以放射性同位素 ^{32}p 或非放射性（酶、地高辛等）标记的人工合成序列特异寡核苷酸（SSO）探针进行杂交，从而分辨几乎所有等位基因的特异性，也能分开差异仅仅是 1～2 个碱基对的亚型。此法具有技术简单，可以同时分析多个样品等优点。但是必须进行探针的标记，目前在国外多采用荧光

标记法。PCR - SSO 是目前应用最多的一种简单、快速而又精确的 HLA - Ⅱ类抗原分型方法，能鉴定所有已知序列的 HLA - DR、DQ、DP 等位基因，精确地分析 DNA 的多态性。

除了上述将扩增产物先固定在膜上，用标记好的已知序列的寡核苷酸探针去杂交检测的“正向杂交”外，也可以进行“反向杂交”。所谓的“反向杂交”是先将已知序列的寡核苷酸按顺序固定在膜上，在 PCR 扩增过程中标记扩增产物或在扩增后标记扩增产物，用这种带标记的扩增产物与固定在膜上的寡核苷酸序列杂交来检测扩增产物的序列，根据寡核苷酸在膜上固定的顺序来确定待测等位基因的序列。正向、反向杂交技术各有特点（表 7 - 6），正向杂交更适合于大批量筛选，反向杂交对较小数量分析更能体现它的优点，目前反向杂交应用较广。

表 7 - 6　正、反向 PCR - SSO 的比较

	正向 PCR - SSO	反向 PCR - SSO
固定在膜上的 DNA	待分析的 PCR 扩增产物	已知序列的寡核苷酸
一次杂交结果	检测多个样品的一类碱基序列差异	一个样品与所有寡核苷酸探针的互补性差异，即一个样品的几类碱基序列差异性
检测 N 个碱基序列差异	需要 N 次杂交	一次杂交
适合性	大批量	小量、大量
寡核苷酸探针要求	除特异性外，无特殊要求	除特异性外，还需要序列长度尽可能一样长
工作量	大	小
操作难度	小	大

2. PCR - SSO 分析中注意事项

（1）PCR 扩增产物的点膜量应尽量调整均匀，不同样品加样量要尽量一致，可以在点量前用琼脂糖凝胶电泳检测一下扩增产物的量。

（2）在同一张膜上设有与杂交探针相应的阳性对照和阴性对照。

（3）在反向杂交膜上应有一个可以判定杂交结果是否为阳性或阴性的控制探针，其序列与所有样品有同源性，没有特异性。只要有扩增产物、杂交条件合适，样品一定显现该杂交点的阳性结果。

（4）杂交时要严格控制杂交温度，如果杂交温度偏低，容易出现非特异性信号；杂交温度偏高，探针不与目的 DNA 片段杂交结合，信号弱或没有。

（5）洗膜时，注意洗膜液的离子强度，严格控制洗膜温度，如进行放射性标记的，随时监测膜的放射性信号。

（6）目的等位基因或等位基因组里的扩增必须是特异性的。如果存在序列类似的其他功能基因或假基因时，必须改变引物序列，以达到特异扩增的目的。

（7）为了避免由于 SSO 碱基长度不同而带来的杂交条件的多样化所产生的麻烦，尽可能地统一寡核苷酸探针碱基长度。对于含有 18 个碱基的探针，选用 54 ~ 55℃杂交温度和

58～59℃洗膜温度。在杂交液中加入 3mol/L TMAC，使杂交的温度只与探针长度有关，碱基序列不影响杂交温度与特异性。

SSO 探针只能基于已发现的等位基因序列而设计，对尚未发现的新的多态性，PCR－SSO 技术虽然可以通过其新的杂交格局来发现，但点杂交有假阳性，当一个样本在某个位点只检出某个等位基因时，不易确定其为纯合基因型还是含一个空白基因的杂合子。

7.4.2 SSP－PCR 技术

序列特异引物－PCR（sequence specific primer PCR，SSP－PCR）（或 PCR－SSP）：是根据基因座某一碱基的差异设计一系列引物，在确定某一碱基为该等位基因所特有的基础上设计一对引物，两引物 3′端的第一个碱基均与等位基因特异碱基互补，特异性引物仅扩增与其相应的等位基因，而不扩增其他的等位基因（图 7－5）。因此，PCR 扩增产物的有无是鉴定特异性等位基因的基础，这种特异性的 PCR 扩增产物可通过琼脂糖凝胶电泳检出。该方法操作简单快速，实验结果容易判断，杂合子也很易于检出，已应用于 HLA－DRB 的分型中。不足之处在于，为检出所有的等位基因，必须用多个引物进行扩增。

等位基因特异 PCR（allelic specific PCR，AS－PCR）原理与 PCR－SSP 一致，如血型 MN 基因型的分析，就是利用 AS－PCR 技术。

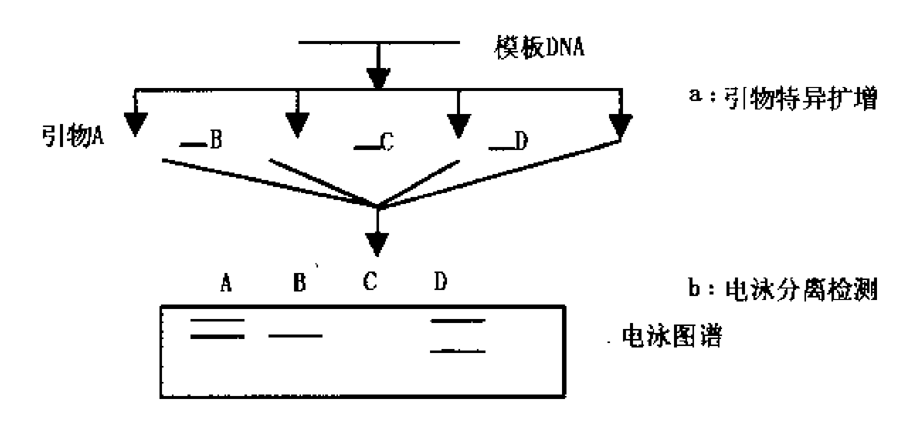


图 7－5　SSP－PCR 原理图

a：用各种不同序列引物分别在不同反应管中扩增，图中 A、B、C、D 代表四种不同序列引物；b：A、B、C、D 四种引物扩增产物加样在同一凝胶不同泳道中电泳检测。由于模板 DNA 中存在有与 A、B、D 引物退火结合位点，得到 PCR 产物，而 C 引物由于模板中没有其互补序列可以结合，因此得不到扩增，在电泳图谱中没有谱带显示。

7.4.3 PCR－RFLP 技术

PCR－RFLP（restriction fragment length polymorphism，RFLP）是在 RFLP 分析方法的基础上发展建立的一种更为简便的 DNA 分型技术。PCR－RFLP 的基本原理是用 PCR 扩增目的 DNA，扩增产物再用特异性内切酶消化切割成不同大小片段，直接在凝胶电泳上分辨（图 7－6）。不同等位基因的限制性酶切位点分布不同，产生不同长度的 DNA 片段条带。此项技术大大提高了目的 DNA 的含量和相对特异性，而且方法简便，分型时间短。已广泛应用于 ABO、HLA、线粒体 DNA 等序列多态性分析中。

PCR－RFLP 与 DNA－RFLP 分型方法相比，虽有很大的改进，但酶切条件要求较高，受影响因素较多，有时 PCR 扩增产物不能完全被内切酶消化切割；有时区分所有的等位

基因比较困难。

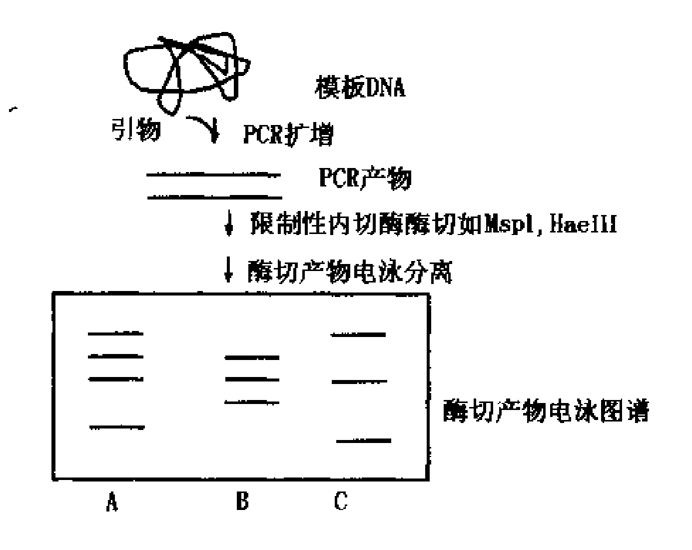


图 7-6 PCR-RFLP 原理图

A、B、C 分别代表三个样品。A、B 样品均有酶切点，C 样品没有酶切点，不能被酶切。

7.4.4 PCR-SSCP 技术

单链构象多态性（single strand conformation polymorphism，SSCP）-聚合酶链反应（PCR-SSCP）是在完成目的 DNA 的 PCR 扩增后进行单链 DNA 多态性分析的一种技术。其原理是在不含变性剂的中性聚丙烯酰胺凝胶电泳时，单链 DNA 因碱基顺序不同所形成的构象不同，电泳迁移率也不相同（图 7-7）。通过 PCR 扩增包括发生单个碱基置换部位及两侧 DNA 片段，变性后进行 SSCP 分析，从理论上可分辨出单个碱基的差异，有效地检出点突变和 DNA 的多态性，有利于探测新的等位基因。

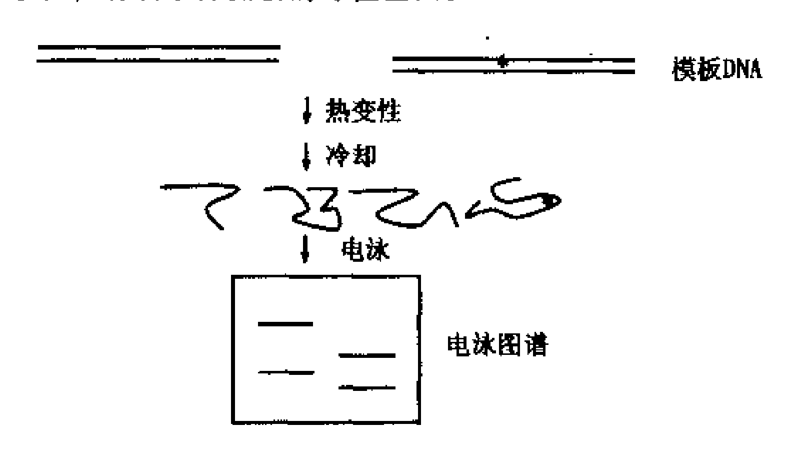


图 7-7 SSCP-PCR 原理图

* 表示碱基差异处

1. 方法

PCR-SSCP 法的基本程序为：

(1) 目的片段扩增　首先以目的 DNA 为模板，采用同位素或非同位素标记的引物或

dNTP 进行 PCR，扩增出目的片段。

（2）变性扩增产物 扩增产物加变性上样液（95%甲酰胺、20mmol/L EDTA，0.03%二甲苯菁蓝、0.05%溴酚蓝），95℃加热使双链 PCR 产物变性，迅速放入冰水中 3min。

（3）电泳 将变性好的扩增产物加在非变性 PAG 上电泳，在电泳过程中，单链 DNA 根据它们的一级结构折叠呈三维形状，电泳迁移率成为单链分子形状的参数。

（4）结果显现与阅读 根据 DNA 片段是否带有标记物，带有什么样的标记物进行显现检测结果。目前主要有放射自显影或银染或荧光标记的 SSCP 方法。

SSCP 适用于检测 400bp 以下片段长度的多态性，因此较长片段在电泳前应用限制性内切酶处理切成较小片段后电泳。

2. 影响因素

PCR－SSCP 检测突变的能力依赖于突变对分子折叠的影响以及分子折叠对电泳迁率的影响。目前还不能在理论上预测这些因素，因此在理论上估计 PCR－SSCP 的灵敏度非常困难，根据经验估计其灵敏度也很困难，因为这种灵敏度明显依赖于被检测片段的序列。目前所能得到的数据还很有限，不能对所有可能的序列预测其 PCR－SSCP 分析的灵敏度。根据现有的经验，对 200 个核苷酸长的片段进行 PCR－SSCP 分析，能检测出超过 90%的所有单碱基替换，对于 400 个核苷酸长的片段能检测出超过 80%的所有单碱基替换。

检测片段越长，PCR－SSCP 分析的灵敏度越低，因此要把长的片段先降解成短的片段再进行 SSCP 分析。

SSCP 方法分析的 DNA 常常是只有一个碱基的序列差异，要求电泳分辨率很高，那么对电泳的要求就高。以下一些因素影响 SSCP 分析：

（1）PCR 扩增 PCR 扩增必须是特异的，没有非特异带。如果 PCR 扩增条件不是特异的，在 PCR 产物中除了目的片段外，还含有非特异的长度可能不同的片段，结果在 SSCP 图谱上会出现许多迁移率各不相同的谱带，使结果无法解释。

（2）变性 DNA 链必须完全变性，变性效率是一个重要参数。如果 DNA 不完全变性，存在部分折叠或退火成双链，均能降低单链 DNA 量，或者导致部分折叠，使结果解释困难。甲酰胺变性效果最好，氢氧化甲基汞（methy mercuric hydroxide）也可做作变性剂。

（3）电泳条件（缓冲液组成、胶基质选择） 没有合适的预测单链 DNA 三维构象的理论模式，即使三维构象可以预测，电泳过程中分离迁移率也不同。下列因素影响 SSCP 分析灵敏度：

①凝胶基质 最好的是聚丙烯酰胺，其孔径小，适合差异 DNA 链分离。用低浓度的聚丙烯酰胺凝胶，凝胶的自由度和脆性增加，滤过孔径增大，对构象改变敏感性增加；采用低浓度的交联剂凝胶，可更好地区分互补的单链（图 7－8）。一般采用 8% ～10%T，1.3%～2.6%C 效果最好。而且采用较薄的凝胶（如 0.3～0.4mm），加快散热。

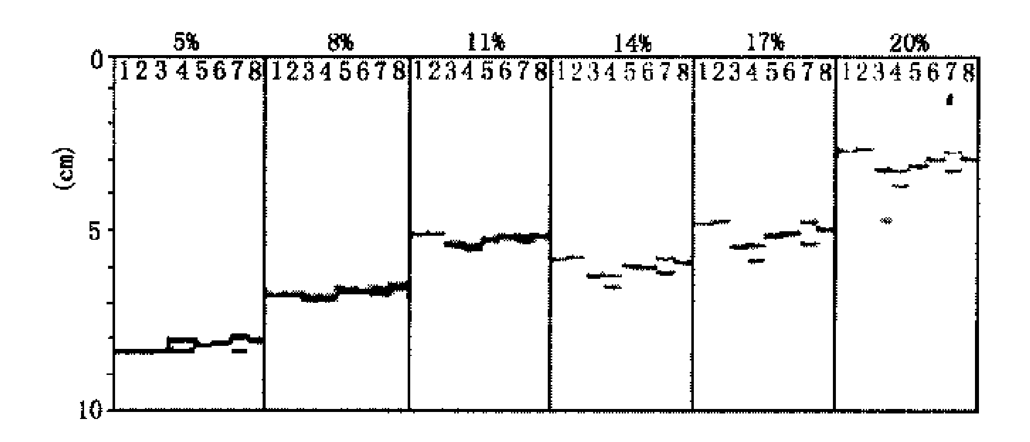


图 7-8 凝胶浓度对分离结果的影响。11%-20%的效果比较稳定。
1: 0101, 2: 0102, 3: 0103, 4: 0103/0201, 5: 0301, 6: 0401, 7: 0501/0101, 8: 0601

②DNA 片段大小 片段长度在 135~155bp 时敏感性最高，大于 200bp 而小于 350bp，敏感性降低 80%；碱基种类及突变位置不影响敏感性。对于大于 1kb 以上片段，先用限制酶切成小片段，这样不降低敏感性。

③缓冲液组分 包括 pH、离子强度，一般采用 0.5×TBE 或 1×TBE。

④缓冲液添加物 凝胶中加 10%~15%蔗糖或甘油等中性剂，或加低浓度的变性剂可以增加敏感性，如 5%尿素或甲酰胺，可以提高敏感性。其机理不清，推测可能轻微改变 ssDNA 的构象，部分打开 ssDNA 分子的卷曲构象，增加分子表面积，降低 ssDNA 在聚丙烯酰胺凝胶中的迁移率，也可能是甘油本身的粘滞性作用。

⑤电泳中凝胶温度 电泳过程中温度对分辨率影响很大，保持恒温对 SSCP 分析结果的重复性非常重要。使用薄的凝胶（如 0.3mm）和高效空气冷却或水套冷却系统以充分散发热量。不加甘油，4~10℃条件下电泳一般总能得到较高的分辨率。在凝胶中添加甘油至浓度为 5%~10%，在 20~25℃条件下电泳也能得到较满意的结果（图 7-9）。

⑥DNA 浓度 DNA 量太多，降低分辨率。

上述任何因素改变会增加或降低分辨能力。经验证明，开始应用 SSCP 时选择室温（20~25℃）加甘油或不加甘油，或 4℃下不加甘油，并且采用两种胶浓度进行分析。

3.SSCP-PCR 技术特点

SSCP-PCR 技术简便，具有较高灵敏度，不需要特殊仪器设备，可以不用特殊放射标记等，是一种简单有效的检测变异基因的方法。但存在一些问题：

(1) 分析片段长度受到限制，为获得理想敏感性，DNA 片段要求在 400bp 以下；

(2) 缺乏完整的理论基础来预测 DNA 片段的迁移率和序列对应构象的关系；

(3) 影响因素多，需要分析人员具有丰富的经验；

(4) 凝胶电泳结果有时难以解释。

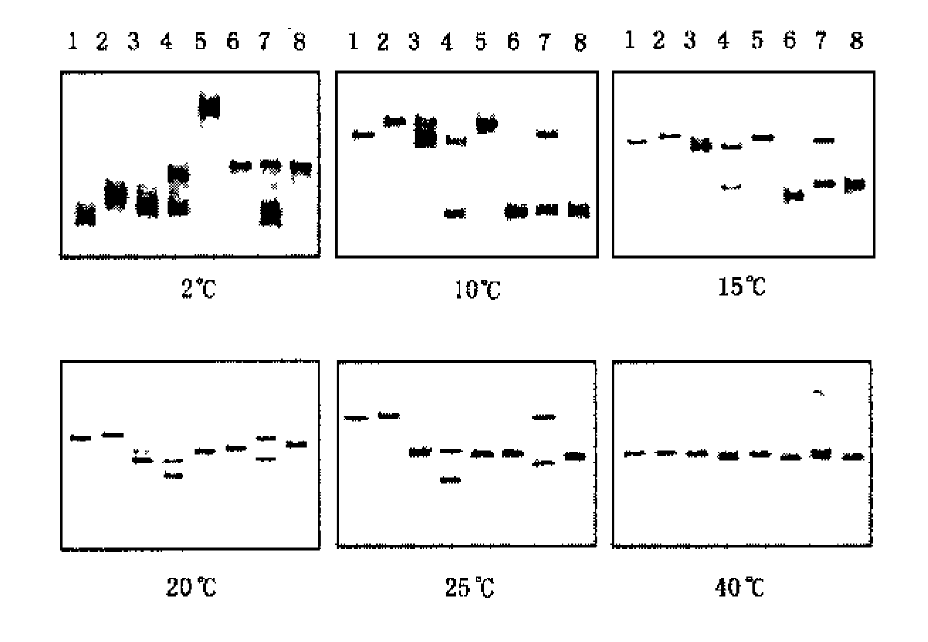


图 7-9　温度对对 SSCP 的影响

HLA-DQA1 的型判定结果 Line 1：0101，2：0102，3：0103，4：0103/0201，5：0301，6：0401，7：0501/0101，8：0601

虽然越来越多的证据表明 PCR-SSCP 分析灵敏度很高，但是并不能排除得出错误结果的可能性。在设定的电泳条件下，一些突变检测不到，不能证明没有发生过突变。

7.4.5 MVR-PCR 技术

MVR-PCR（minisatellite variant repeat mapping by PCR），即小卫星重复单位多态性分析。有些小卫星，除了重复单位串联数目有差异外，重复单位内部还存在序列变异，例如 D1S8 基因座，大约有一半重复单位内部含有 A→G 碱基替换，使这类发生替换的重复序列内出现了 HaeⅢ 酶切点。根据碱基差异设计分别能扩增两种重复单位的下游引物，两引物 32-TAG-T 和 32-TAG-A 序列差异只在 3′端一个碱基差异，基因座上游 5′端为共同的引物 32D。32-TAG-A 专门扩增含 HaeⅢ 酶切点的重复序列，32-TAG-T 专门扩增不含 HaeⅢ 酶切点的序列。两个扩增体系均扩增出一系列不同长度的片段。32-TAG-A 扩增片段为 a 型，32-TAG-T 则为 t 型片段。两种扩增产物经凝胶电泳分离检测，按片段长度依次排列，比较两泳道相应位置有或无片段进行分型。如果只有 a 片段而无 t 片段编码为 1（aa 型），有 t 无 a 的为 2（tt 型），既有 a 又有 t 片段的为 3（at 型），最后形成一系列数字表示的编码来显示个体的差异（图 7-10）。MVR-PCR 只考虑两组扩增产物电泳条件的一致性，而不涉及片段长度的测定。MVR-PCR 最大的特点是高度的多态性，一个基因座分析理论上可以进行个体同一认定，而且数字化信息利于计算机管理与保存。

7.4.6 RAPD 技术

随机引物扩增 DNA（random primer amplified polymorphic DNA，RAPD 或 arbitrary PCR，A-PCR）的原理是利用 10bp 长的一条链引物进行 PCR 扩增。引物的序列是随机的，在基因组 DNA（模板 DNA）中有许多结合位点，扩增出几个甚至几十个大小不同 DNA 片段

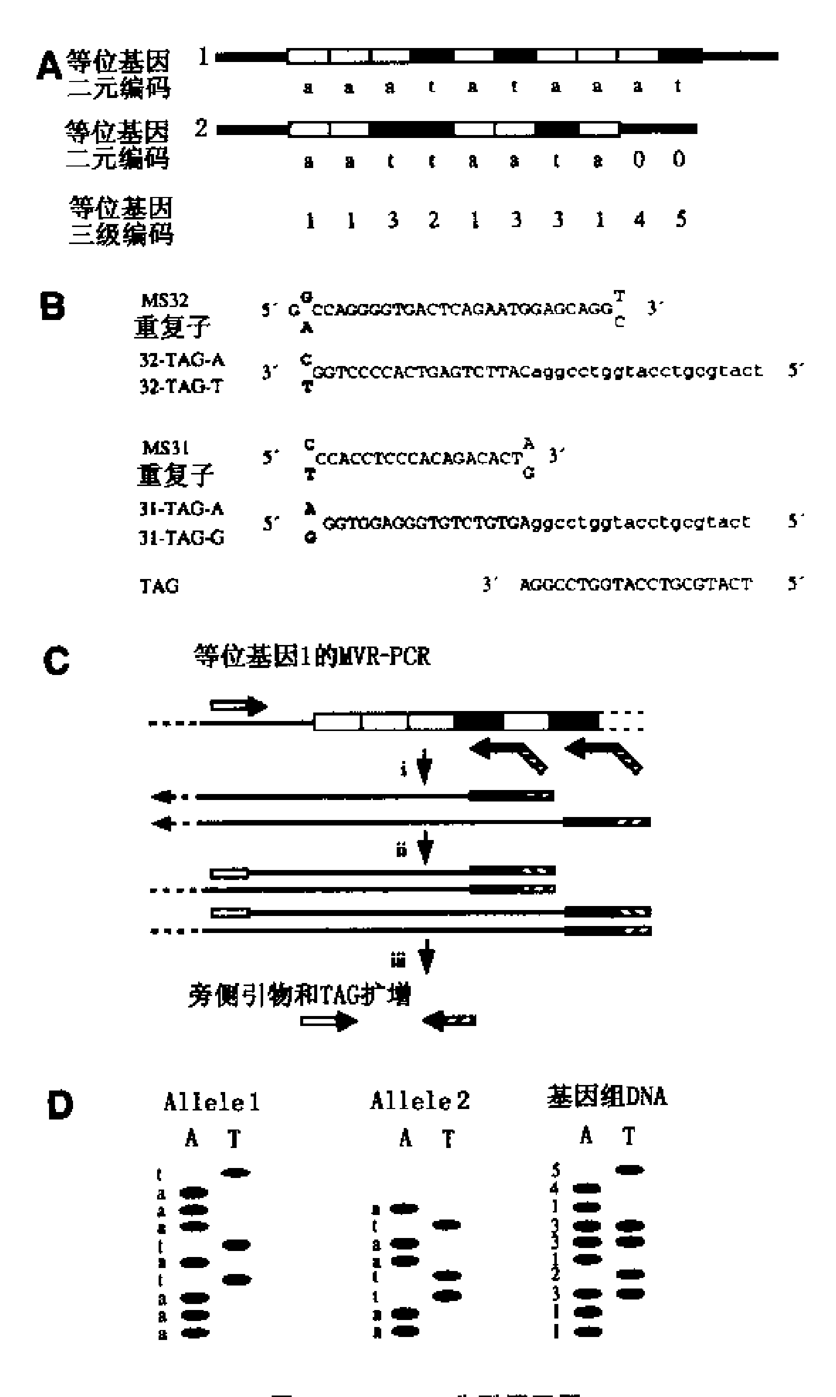


图 7－10 MVR 分型原理图

A：等位基因重复单位；B：重复单位序列；C：以 t 型片段为例扩增；D：MVR 编码

(图 7－11)。引物设计时无需了解靶 DNA 侧翼序列，可以分析基因组中尚不了解 DNA 多态性基因座和等位基因片段。随机引物 PCR 检测的信息量多少，取决于引物的序列与长度。随机引物扩增方法可用来增加基因组 DNA 模板量，它在动物和植物品种筛选中应用较多。由于其重复性问题，在法医学多态性分析中实际应用的比较少。

7.4.7 巢式 PCR 技术

巢式 PCR（nested－PCR）或叫嵌合 PCR，是进行二次 PCR 扩增，第二次 PCR 以第一次

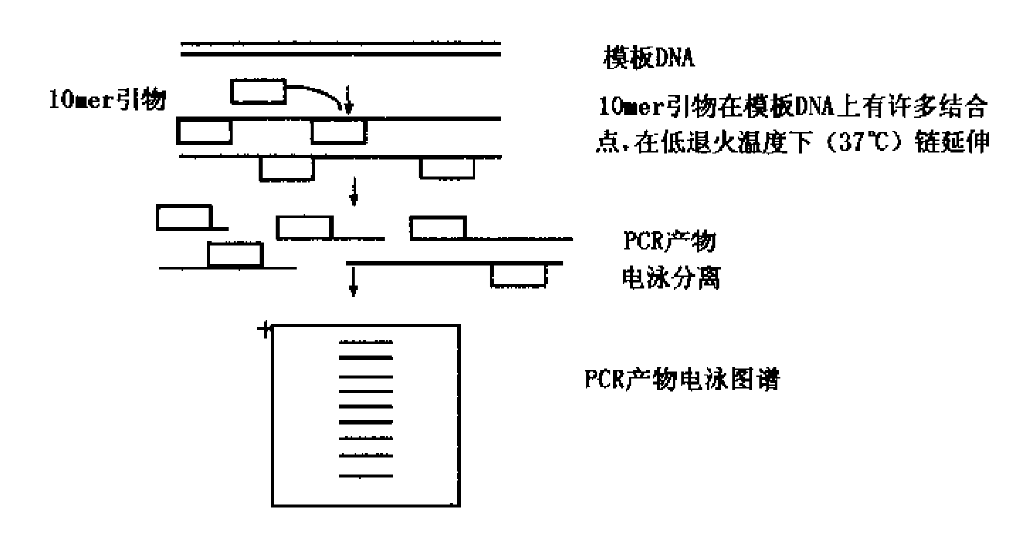


图 7-11 RAPD 原理

PCR 产物为模板，选用第一轮 PCR 引物内侧的序列为引物进行 PCR 扩增，第一次 PCR 扩增产物中部分序列得到第二次扩增（图 7-12），明显提高了检测灵敏度。如果第二次扩增的二个引物均在第一次 PCR 引物序列内，称之为巢式 PCR，如果有一条引物与原来的引物相同，只有一条在内侧的 PCR 称之为半巢式 PCR（semi-nested PCR）。巢式 PCR 主要用来提高扩增灵敏度和特异性，如线粒体测序反应中测序片段制备就是进行的巢式 PCR。

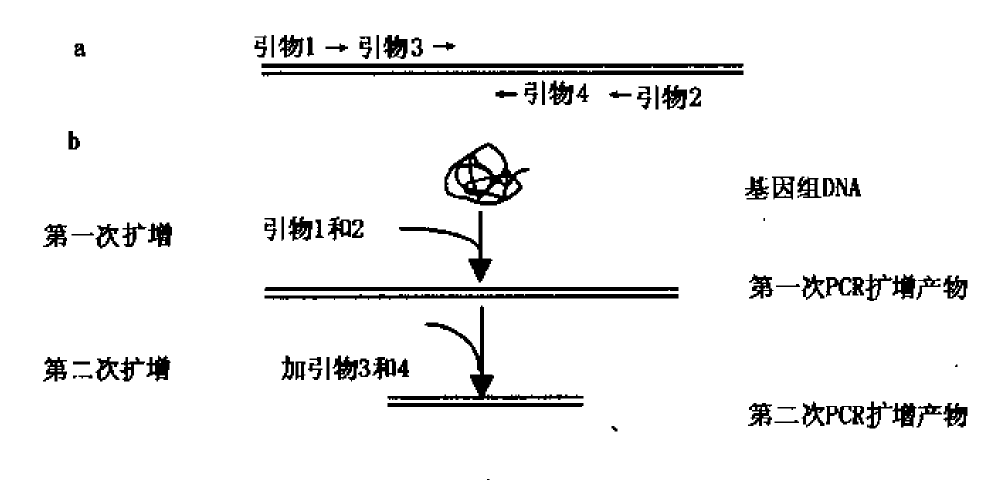


图 7-12 巢式 PCR 原理图

a：巢式 PCR 引物结合区　　b：巢式 PCR 扩增

7.4.8 RT-PCR 技术

RT-PCR 即逆转录 PCR（reverse transcriptase polymerase chain reaction，RT-PCR）是 RNA 的快速体外扩增，原理是 RNA 先通过逆转录酶合成一条 cDNA，再以 cDNA 为模板进行扩增。RT-PCR 主要用来检测 RNA 病毒、含 RNA 的微生物、检测或比较某一 mRNA 的表达。有报道，应用 RT-PCR 进行损伤研究。

7.5 防止 PCR 污染的对策

由于 PCR 能够扩增甚至 1 个分子的 DNA，极微量的污染便可导致假阳性结果，所以应当注意防止反应体系被痕量 DNA 所污染。PCR 中污染源有四个：(1) 标本间的交叉污染；(2) 实验室中非人 DNA 污染；(3) 技术人员的 DNA 污染（如头皮屑、脱落细胞和毛发）；(4) PCR 产物的污染，也叫"残留污染"，其中以最后一种污染源最严重。PCR 产物一般为 10^{13} 拷贝/ml，取 0.1μl 即为 10^{9} 拷贝。因此，尤其是在待扩增靶序列浓度低的情况下，更要采取防范措施。

7.5.1 污染的预防

扩增中最可能造成交叉污染的方式是由产物或标本的气溶胶带来的污染（空气污染）。一般来讲，在标本的制备、旋涡振荡混合、打开反应离心管、加样及取下吸头等过程中，都可能产生这种气溶胶。因此在 PCR 操作过程中一定要避免或减少这种气溶胶的生成。进行 PCR 时，按照下述规则可有效地预防 PCR 的污染。

1. 隔离不同操作区

分隔操作区域是最重要的预防措施，PCR 操作区可划分为标本处理区、反应混合液准备区和 PCR 扩增区以及产物分析区。PCR 反应准备应在装有紫外灯的超净台内进行，除了操作分区外，还应注意操作器材的专用。

2. 分装试剂

PCR 用的去离子水和缓冲液均应高压处理。必须用无菌去离子水配制。所有这些试剂均应在无 PCR 扩增产物区制备，玻璃或塑料用具和移液器等要求是新的、从未接触过 DNA，配好的试剂分装成小体积保存，分装后每管要标记批次、配制时间，以便在发生污染时，比较容易查找原因。同样，引物也应在无 PCR 扩增产物环境中合成与纯化。

3. 严格实验操作

除了残留污染是造成假阳性的重要原因外，样品间的交叉污染也是不容忽视的原因之一。因此，不仅要在进行扩增反应时谨慎，在样品处理的所有环节中都应注意，下面几点应格外注意：

(1) 戴一次性手套操作，进出不同隔离区时都要换手套；

(2) 避免反应液的飞溅。开盖时尤应注意，在开盖前将试剂管离心几秒使壁上和盖上的反应液集于管底。一旦发生飞溅，应立即换手套；

(3) 用一次性移液器或吸头。吸头不要暴露于空气，防止产物气溶胶的污染。加样器最易受产物 DNA 或标本 DNA 气溶胶的污染，所以最好使用可替换的加样器，使用带滤膜的吸头。如由 Raini (Microman) 生产的加样器，既具有一次性吸头，也有一次性插人配件。这些器件是完全分隔的。因此，不会产生上述交叉污染。应经常擦拭加样器与吸头接触部位。如有条件最好经常更换加样器以防 DNA 的污染和确保加样的准确性；

(4) 配制总反应混合液，先将可混匀的成分（dNTP，缓冲液，引物和酶）混合后再分装不同反应管。这样，可以减少操作，避免污染，增加反应的精度；

(5) 样品制备应按无菌操作的原则进行，避免样品间的相互污染；

(6) 最后加样品 DNA。加人 DNA 后应盖紧反应管。

7.5.2 **污染源的追踪**

1. 设立阳性与阴性（空白）对照进行监测

设置阴阳性对照进行监测，监测反应体系各成分的污染情况。阳性对照选择已知基因型、且重复性好的样品。试剂空白对照应包括 PCR 所需的全部成分，仅不加模板 DNA，如果出现扩增产物，表示试剂已有污染。

2. 检查环境污染源

除了上述试剂污染，易引起污染的还有环境污染。包括：①离心管与 tip 头；②加样器；③离心机；④超净台；⑤紫外灯箱；⑥干冰/乙醇或水浴锅；⑦冰箱门把手、冷冻架、门把手和实验台面等。此时可用擦拭试验（wipe test）来追踪可疑的环境，方法是：①用无菌水浸泡的灭菌棉签擦拭可疑污染源；②将擦拭过的棉签浸泡在 0.1ml 去离子水中；③取 5μl 浸泡液做 PCR 实验；④电泳检查结果，观察是否有 PCR 产物，若有 PCR 产物，说明有污染的存在。

7.5.3 **污染的处理**

1. 环境污染

（1）稀酸处理法：对擦拭试验阳性部位或器件可用 1mol/L 盐酸擦拭或浸泡，使残存 DNA 脱嘌呤。

（2）紫外线照射法：UV 波长一般选择 254～300nm，照射 30min 即可。需注意的是，选择 UV 来消除 PCR 产物残留污染时，要考虑 PCR 产物的长度与产物序列中碱基的分布。UV 只对 500bp 以上长片段有效，UV 照射时，PCR 产物中嘧啶碱基会形成二聚体，形成二聚体的程度取决于 UV 波长，这些二聚体可终止延伸。

2. 反应液污染的处理

反应液污染时，可采用下列方法之一处理：

（1）DNase Ⅰ法：在不加模板和 Taq 聚合酶的 PCR 混合液中加 0.5u DNase Ⅰ，室温反应 30min 后，加热灭活，然后加入模板和 Taq 聚合酶，进行正常 PCR。

（2）内切酶法：选择识别 4 个碱基的内切酶（如 Msp Ⅰ等），进行同 DNase Ⅰ法反应处理，作用时间延长至 1h。最好几种酶联合使用，可克服用一种酶识别序列特异的缺陷。

（3）UV 照射法：反应液中不加模板与 Taq DNA 聚合酶，注意事项与方法同上述 UV 照射法。

（4）γ线辐射法：γ线辐射能通过水的离子化产生自由基破坏 DNA。1.5kGy 的辐射可完全破坏 0.1ng 基因组 DNA，2.0kGy 可破坏 10^4 拷贝的质粒分子；4.0kGy 的辐射不影响 PCR，但高于此限时会使 PCR 效率下降。引物受照射而不影响 PCR。

3. 尿嘧啶 DNA 糖基化酶（UDG）法

（1）原理：这种方法利用 UDG 可裂解单或双链 DNA 中的尿嘧啶碱基和糖磷酸骨架间的 N－糖苷键，使含 dU 的 PCR 产物降解，阻止 Taq DNA 聚合酶的延伸，但 UDG 对不含 dU 的模板无任何影响。在 PCR 产物或引物中用 dU 代 dT，这种 dU 化的 PCR 产物与 UDG 一起孵育后使其失去再被扩增的能力。UDG 对 RNA 中的尿嘧啶和单一尿嘧啶分子则无任何作用。

（2）dUTP 法：用 dUTP 代替 PCR 中的 dTTP，使产物中掺入大量 dU。在 PCR 正常循环

前，用 UDG 在 50～55℃处理 PCR 混合液可以消除 PCR 产物的残留污染，在正常 PCR 循环前进行 95℃处理 5min 灭活 UDG，含 dU 的新 PCR 产物不会受到影响。

（3）dU 引物法：合成引物时以 dU 代 dT，PCR 产物 5′端带 dU。UDG 处理后，引物失去了结合位点而不能扩增。对 1～2kb 以上的长片段扩增，用 dUTP 法的扩增效率比用 dTTP 的低，dU 引物则没有这一现象。dU 引物最好将 dU 设计在 3′端或近 3′端。

UDG 法的优点：①去除任何来源的污染；②UDG 处理与 PCR 扩增可在同一试管内进行；③污染消除彻底。

7.6 PCR 常见问题及处置

1. 没有扩增产物形成：①循环温度，特别是变性温度是否正确，注意 PCR 仪的指示温度是否与实际温度相符。②引物的组成是否正确，有无二级结构的形成。③反应系统有无蛋白酶或核酸酶的存在，使聚合酶或模板及产物降解，可在 95℃以上处理反应系统使其灭活。④某些来源的 DNA 可能有较难去除的蛋白质、色素等成分，抑制 PCR 系统，用蛋白酶 K 重新处理，或反应体系中添加 BSA。⑤Taq DNA 聚合酶活性下降或失活，要更换新酶。⑥DNA 模板量太低或已降解，可以用琼脂糖凝胶电泳检查一下模板 DNA。

2. 非特异性产物的出现：①提高退火温度，缩短退火或延伸时间；②调整引物、Taq DNA 聚合酶或模板用量；③调整 Mg^{2+} 用量；④减少热循环次数。

3. 引物二聚体的形成：①检查引物的序列，特别是在 3′端是否有互补区；②提高退火温度；③调整引物与模板浓度，使模板比例增高；④增加引物长度，减少热循环次数。

第八章　DNA 序列测定

8.1 概述

DNA 测序，即对 DNA 的一级结构——碱基序列进行测定。1965 年 Holley 用经典法测定了酵母 tRNA 和 75 个核苷酸顺序，随后又陆续发展了指纹图图谱法、加减法、化学直读法等测序方法。1977 年 Sanger 创立了测定 DNA 序列的双脱氧核苷酸（ddNTP）链终止法，同年 Maxam 和 Gilbert 创立了另一种 DNA 测序方法——化学降解法。前者是通过 DNA 多聚酶以单链 DNA 为模板合成一系列长短不同的 DNA 片段，后者采用特殊化学试剂将 DNA 链进行化学降解产生一系列长短不同的 DNA 片段。两种方法虽原理不同，但都生成互相独立的若干组带放射性标记的寡核苷酸，每组寡核苷酸都有固定的起点，但随机终止于特定的一种或更多种核苷酸上。这些寡核苷酸产物的长度由某一特定碱基在原 DNA 片段中的位置所决定。然后把几组寡核苷酸加样在测序凝胶中电泳分离，从凝胶的放射性胶图上直接阅读 DNA 序列。此法可区分长度仅差一个核苷酸的 DNA 片段。

在链终止法和化学降解法的基础上，不断发展，测序模板制备、测序反应策略及测序反应产物检测等方面都有了很大变化，逐渐形成越来越完善的全自动序列测定系统。从早期通过分子克隆技术将目的 DNA 插入单链线状噬菌体 M13 载体上，提取单链模板用于序列测定，到质粒双链 DNA 进行测序，以及 PCR 技术发展、应用，建立了更为简单、方便的直接对体外扩增 DNA 产物进行序列测定方法；在测序反应策略上经历了 DNA 聚合酶 I Klenow 片段测序、高温 DNA 聚合酶（Taq DNA 聚合酶）、测序酶及测序酶 2.0 版测序，直至目前广为应用的 PCR 测序——循环测序法。还有应用在测序反应产物检测上的其他领域的各项技术，如垂直凝胶电泳结合同位素标记、非同位素标记、引物四色荧光标记及末端荧光标记检测，单道毛细管凝胶电泳以及最近的多道毛细管分析使 DNA 测序自动化程度和效率更高。还有用于序列比较目的的固相微型测序技术。

本章主要介绍链终止测序法和化学降解 DNA 测序法原理及技术，以及各有侧重地介绍后来发展起来的非同位素测序、银染测序、荧光测序、循环测序和固相微型测序等。

8.2 测序策略

测序方法很多，实际应用中采用何种方法，主要取决于实验室条件和测序的目的。测序的目的有：①测定未知序列或叫从头测序；②确定重组 DNA 的方向和结构；③对突变进行定位和鉴定；④比较研究。后三种是确证性测序，法医学中测序主要就是这种测序。未知序列测定是指提供一个未知序列的全长、准确的核苷酸序列。目前的测序系统一般最多可以阅读 1 000 个核苷酸的序列，单套测序反应所能准确测定的 DNA 序列最长能达到 400～500 碱基。因此，克隆到测序载体上的目的 DNA 长度以 400～500bp 最为适宜。

1. 对于较小片段（如小于 500bp）的目的 DNA

可以直接利用质粒系统（如 pUC18 等）测序。目前其测序水平基本达到 M13 单链 DNA 序列测定的水平。因此，除非特殊用途（如体外定点突变），都可以用 PUC 系统代替 M13 系统。

2. 对于大片段的目的 DNA

可考虑以下几种方法：

(1) 限制性片段亚克隆法　对目的 DNA 进行酶切分析，确定酶切图谱后，以合适的限制酶消化，制备各种长度的限制酶片段，克隆到测序载体上建立亚克隆库。然后测定亚克隆的 DNA 序列，通过排列分析，可确定目的 DNA 片段的全序列。然而实际上应用起来相当不容易，特别是当目的 DNA 片段中限制酶位点分布不均，则上述策略很难奏效。因此，仅当目的 DNA 片段限制位点分布均匀，片段不太大的情况下可考虑采用这一策略。

(2) 定向法　在定向法中，靶 DNA 的测序按计划有秩序进行。例如，利用通用引物测定两端的目的 DNA 序列，然后以已测定的序列为基础来设计新的多聚寡核苷酸作为后续一套反应的引物，从而循序渐进地获得目的 DNA 的全部序列，因此称之为渐进法（图 8－1）。另一种方法是，通过测定一系列嵌套的缺失突变体的序列而获得，这些突变体具有共同的缺失起点（通常在目的 DNA 的一端）并分别插入目的序列不同的位置，即这些突变体缺失序列起始于目的 DNA 一端，逐步延伸到目的序列不同距离。因此可使目的 DNA 中较远距离的片段渐渐地进入通用引物测序的范围之内。

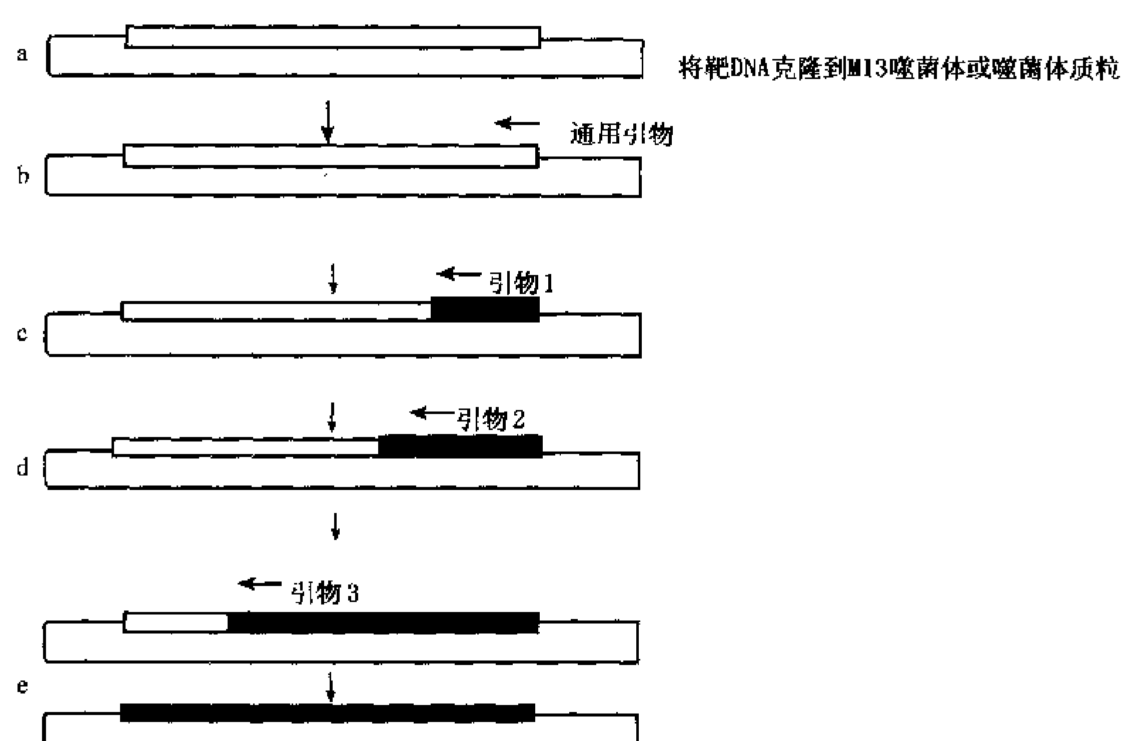


图 8－1　渐进法定向测序

(a) 目的 DNA 片段克隆到质粒载体上；(b) 利用通用引物（正、反向）对克隆片段进行序列测定；(c、d) 以测定的序列为基础设计并合成寡核苷酸作为引物向前测序；(e) 双向测序重叠得到目的 DNA 的全长序列

(3) 随机法(或称鸟枪测序法) 利用超声波或 DNA 酶将目的 DNA 随机打断成大小不同的片段并获得亚克隆，收集这些亚克隆的 DNA 序列资料，通过计算机处理而获得目的 DNA 的全部序列(图 8-2)。这种方法由剑桥的医学研究委员会(M.R.C.)实验室率先推行，曾经成功地用于人线粒体(Anderson 等，1981)、人腺病毒 DNA、λ 噬菌体 DNA 以及 Epstein-Barr 病毒 DNA 的序列测定。采用此法，所要测定的序列通常会比待测 DNA 的实际长度要多 4~6 倍，在大多数情况下，直到双链 90% 左右的序列测出以后，才能得到全长序列。由于用于测序的亚克隆是随机挑选出来的，因此某些区段的序列会在全序列测定之前被重复测定，有时需要很长时间才能确定最后几个亚克隆的序列，而使全长序列得以拼接完全。该策略测序的前提是建立随机重叠克隆文库。

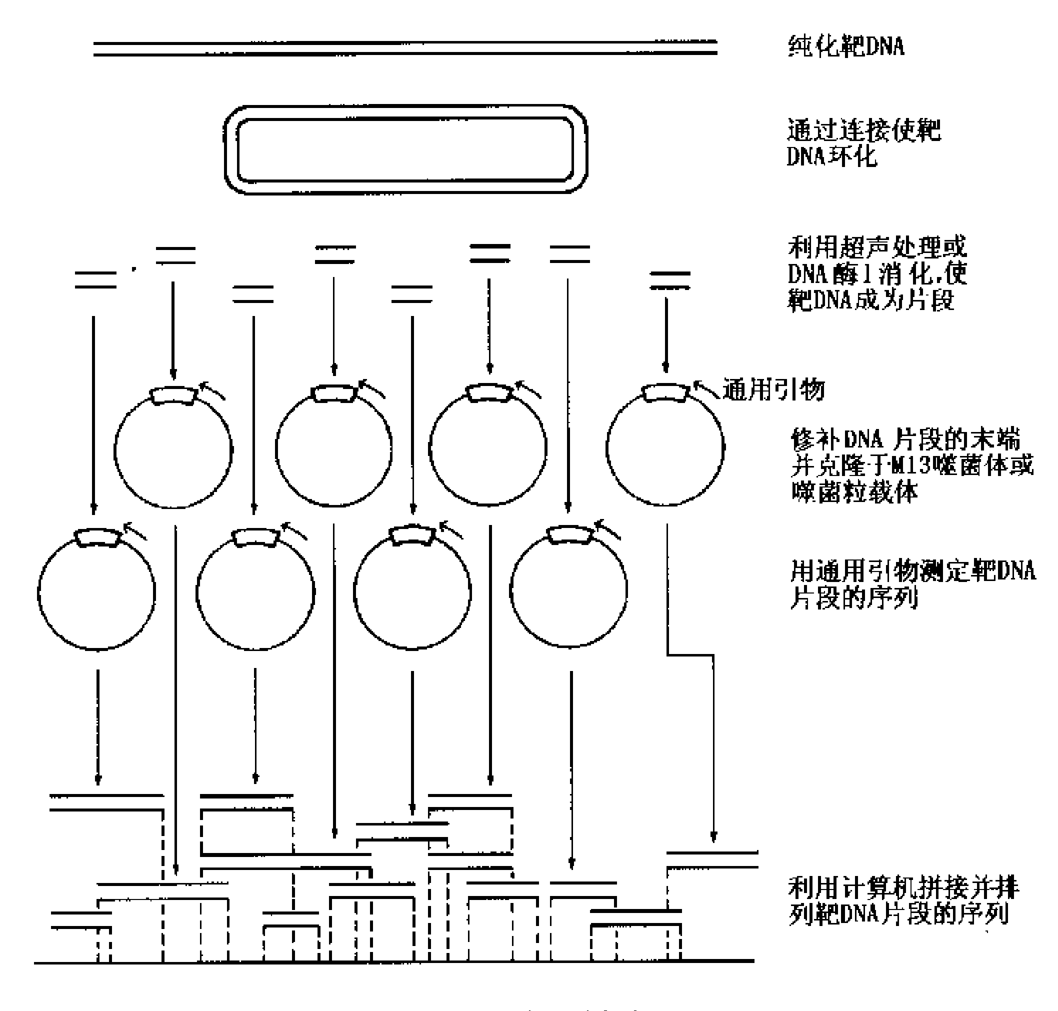


图 8-2 随机测序法

8.3 双脱氧链终止法测序

8.3.1 双脱氧链终止法测序的酶学原理

DNA 的复制有 4 个基本条件，即 DNA 聚合酶、单链 DNA 模板、带有 3'OH 末端的寡核苷酸引物、4 种 dNTP(dATP, dCTP, dGTP, dTTP)。在这一反应体系中，引物与模板退火

形成双链区后，DNA 聚合酶便结合到 DNA 双链区上而启动 DNA 的合成；在引物的引导下，聚合酶在模板链上沿 3'→5'的方向移动，利用体系中 4 种核苷酸聚合成一条与模板链互补的 5'→3'的 DNA 新生链。如果在 DNA 体外合成的反应体系中加入双脱氧核苷三磷酸（2'，3' - ddNTP），DNA 的合成情况则完全不一样。2'，3' - ddNTP 与普通的 dNTP 的不同之处在于它们在脱氧核糖的 3'位置上缺少一个羟基。这些 ddNTP 可以在 DNA 聚合酶作用下，通过 5'三磷酸基团与正在延伸的 DNA 链的末端的脱氧核糖 3' - OH 发生反应，形成 3'，5' - 磷酸二酯键而掺入到 DNA 链中，但它们本身没有 3' - OH，不能同后续的 dNTP 形成磷酸二酯键，正在延伸的 DNA 链在此终止。因此，在 DNA 测序系统中，除了加入正常反应所必需的 4 种 dNTP 外，如果再加入一种一定比例的 ddNTP（例如 ddATP），则新合成的 DNA 链在掺入正常 dNTP（例如 dATP）的位置都有可能掺入 ddNTP（如 ddATP）而导致新合成链在所有的含有该碱基（如 A）位置上终止。由于产生这样一种 ddNTP 与 dNTP 的竞争，反应产物为一系列长度不同的多核苷酸片段。这些片段具有共同的起点——引物的 5'端，但其终点（ddNTP 掺入的位置）不同，其长度取决于 ddNTP 掺入的位置与引物 5'端之间的距离。测序反应时，一般设置一套四组反应（A、T、G 和 C），每一组反应对应一种碱基，体系中除加入的 ddNTP 种类不同外，其余成分相同。如图 8 - 3，A 管中除加入 dATP（^{32}P 标记作为示踪物）、dGTP、dCTP、dTTP 还加入一定浓度的 ddATP，这种反应混合液被称为 A 反应混合液，其他依此类推。在 A，G，C，T 各反应混合液中通过控制 ddNTP 与 dNTP 的浓度，则可获得与模板互补的各种长度的 DNA 新生链、每一种新生链都是被反应混合液 ddNTP 所终止，它们都有共同的起点——引物的 5'端，但终点则由 ddNTP 掺入的位置所决定。再通过测序凝胶电泳分离不同长度的 DNA 片段，特殊的检测方法显示被中止的 DNA 片段长度，确定该位置的碱基种类。目前常用的有同位素标记放射自显影方法、非同位素标记、银染和荧光标记全自动测序。荧光标记全自动测序是用 4 色荧光标记引物（dye - primer）或标记 ddNTP（dye - terminator），使一种荧光与一种碱基对应起来，再经电泳、激光扫描，计算机就能自动读出这个 DNA 的全部核苷酸顺序。

8.3.2 双脱氧链终止测序方法

双脱氧链终止测序方法有三个步骤：（1）测序反应的单链 DNA 模板的准备；（2）链合成反应即测序反应；（3）聚丙烯酰胺凝胶电泳分离片段，读出片段序列。经典的 Sanger 测序反应中，利用 DNA 聚合酶进行一次链延伸和终止反应，需要大量测序模板。而随着 PCR 技术的发现，PCR 技术应用到测序技术，利用 PCR 进行测序模板的制备，也可以利用 PCR 直接测序，即进行循环测序，只需微量测序模板就可以获得理想的测序结果。目前大都采用循环测序方法。不过，无论采用哪一种测序方法，模板的制备是至关重要的。

1. 测序模板准备

两种类型的 DNA 可以用作测序模板，即单链 DNA 和经过热变性或碱变性的双链 DNA。早期的 Sanger 测序应用分子克隆技术将目的 DNA 插入到单链线状噬菌体 M13 载体上，提取单链模板用于序列测定。目前可以应用双链 DNA 进行直接测序，只需将目的片段 DNA 重组到质粒 DNA 中，进行转化，增殖质粒，使模板制备操作更加方便、省时省力。利用 PCR 技术扩增目的片段使模板 DNA 制备不需要克隆。但双链 DNA 的模板质量比 M13 单链 DNA 模板差，正负链同时存在，缺乏模板的一致性。

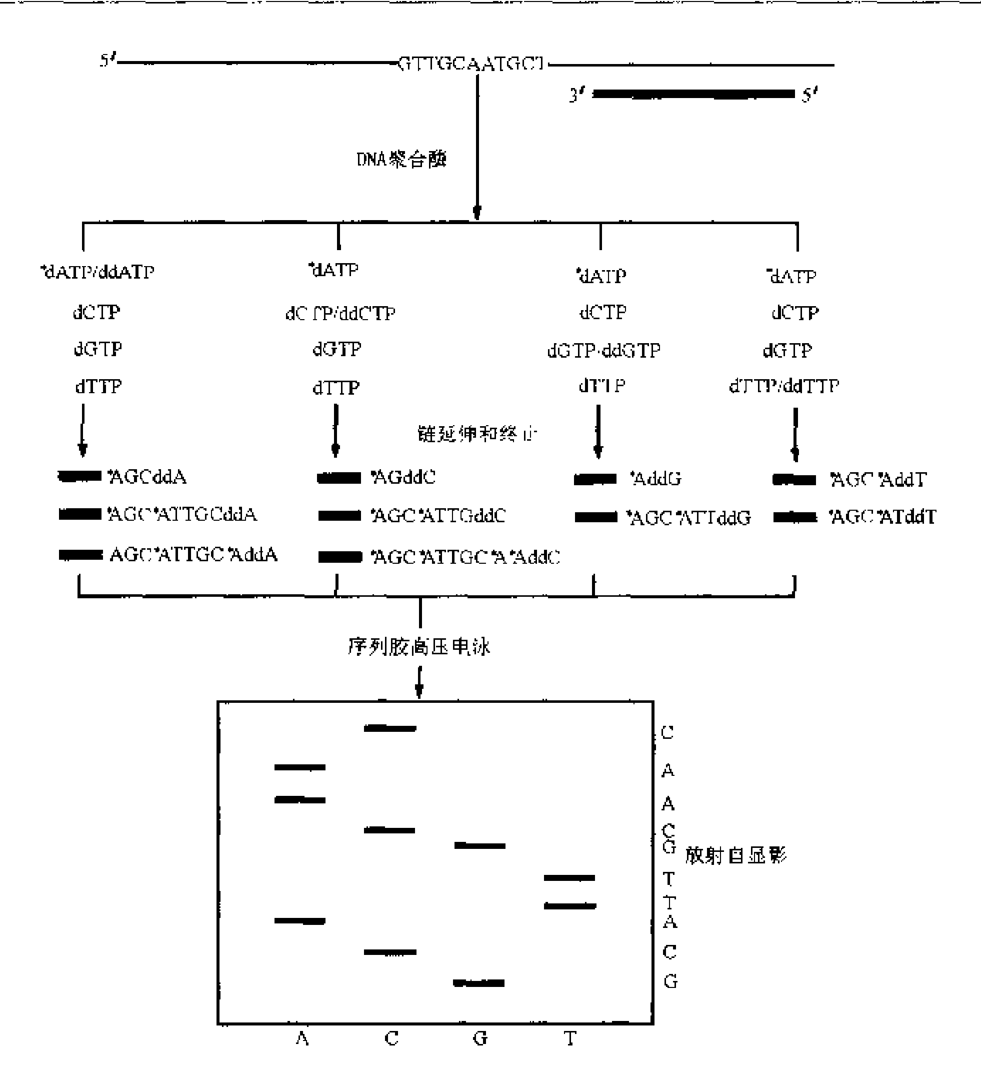


图8-3 双脱氧链终止法测序原理示意图

2. 利用PCR制备测序模板DNA

用PCR方法扩增靶DNA，随后用扩增的DNA产物直接进行测序，已成为分子遗传学分析研究的一项极有用的策略。采用这一策略，可完全省去首先进行PCR产物克隆这一耗时的步骤，而且直接从粗制的生物样品中得到目的DNA序列。靶DNA的拷贝数可少至1到几个的基因组DNA。目前有三种类型PCR方法可以制备测序模板DNA:

(1) 常规PCR产生dsDNA产物，反应结束后通过纯化除去过剩的引物和dNTPs，随后以PCR扩增产物序列互补的一寡核苷酸序列为测序反应引物进行测序分析。

(2) 用带有噬菌体启动子序列的引物进行PCR，然后用噬菌体RNA多聚酶转录获得ssRNA，用反转录酶双脱氧法对ssRNA测序。用反转录酶或Taq DNA聚合酶对其产物直接测序。合成这种带噬菌体启动子序列的引物的成本较高。

(3) 不对称PCR产生单链DNA (ssDNA)，用Taq DNA聚合酶或其他测序酶对ssDNA测序。

标准PCR可以得到双链DNA拷贝，而由Gyllensten和Erlich改进的PCR方法则可以产

生一个单链 DNA。其主要原理是在扩增体系中，利用不同浓度的引物得到 ssDNA。不对称引物的比例一般为 50:1 或 100:1。在开始的 10～15 个循环中，主要产物为 dsDNA。当有限的引物被耗尽，ssDNA 通过过量的引物开始合成，并且随循环次数的增加以线性形式（而非指数形式）增长。这种单链 DNA 作为模板，用原引物或者内部序列作引物测序，内部引物测序可以增加特异性。不对称 PCR 基因扩增的效率比标准 PCR 要低，因为 ssDNA 只是线性增加，而不像常规两条引物退火形成 dsDNA 的倍增。要获得最大量的 ssDNA 需要更多的循环周期，大部分靶序列要求至少 30～40 个周期。

PCR 产物中的 dsDNA 和 ssDNA 的量因引物不同和引物比例不同而变化，最佳的引物比例范围在 0.5:50pmol 和 5:100pmol 之间。另外，进行不对称 PCR 时，dNTPs 的浓度比标准 PCR 低，低浓度容易获得最大量的单链产物，减少某些靶序列的本底产物。测序时，引物退火前不进行高温变性，这样 dsDNA 就不会参与测序反应。

用 PCR 直接从基因组或克隆片段制备测序模板比目前使用的细胞依赖性方法更有效。对于未知序列的克隆片段可用于载体多聚接头或其附近序列互补的寡核苷酸作为扩增引物。用 PCR 制备测序模板可以省去细菌培养、模板提取这些操作。直接对 PCR 双链产物进行序列测定，一个非常关键的步骤是纯化扩增产物，从扩增产物中除去引导 PCR 扩增的寡核苷酸引物（未标记引物）。因为这些未标记的引物与标记引物竞争模板上的结合位点，并产生完全不含标记的延伸产物。通过纯化同时将 dNTP 从 DNA 扩增产物中除去，使这些 dNTP 不干扰测序反应体系中 d/ddNTP 的比例。dNTP 和寡核苷酸可通过在微量浓缩器中进行离心透析或用一些商业 PCR 产物纯化试剂盒纯化予以去除。如果条件不允许，也可采用乙醇沉淀、聚丙烯酰胺和琼脂糖凝胶电泳的方法来纯化扩增产物。

8.4 直接 PCR 测序——循环测序

8.4.1 循环测序原理

在 Murray 方法基础上，已发展出一系列 PCR 介导的测序策略——循环测序策略。它们都是利用热循环仪自动循环的高效能力，使链终止的序列产物获得扩增。一个测序反应循环结束后，测序模板通过高温变性，可以进行链的延伸与终止。每个测序循环由三个步骤组成（图 8－4）。第一步，PCR 扩增的 DNA 经变性成为单链形式；第二步，带或不带标记的引物与其中的一条链上的互补序列退火；第三步，退火后的引物在耐热 DNA 聚合酶催化下发生链延伸－终止反应。由此产生的模板链与延伸终止链形成的部分双链的产物，在下一轮测序循环中，再次被变性，释放出模板链，作为又一轮引发反应的模板，同时积累每轮循环产生的链终止产物。这些循环步骤重复 20～40 次，使链终止产物以线性方式获得扩增。

循环测序，是在对模板进行 PCR 扩增的同时，进行模板与引物的退火、标记、延伸和终止反应。采用耐热 DNA 聚合酶，测序反应可以进行多次循环，因此称之为循环测序，比双脱氧链终止法为基础的标准 DNA 测序在方法学上又有了改进。

8.4.2 循环测序优点

循环测序方法中使用耐热 DNA 聚合酶，应用热循环仪，此法较传统方法有如下优点：

（1）产生清晰可辨的序列梯带所需的模板 DNA 量少，用 P^{32}末端标记引物，4fmol 模板

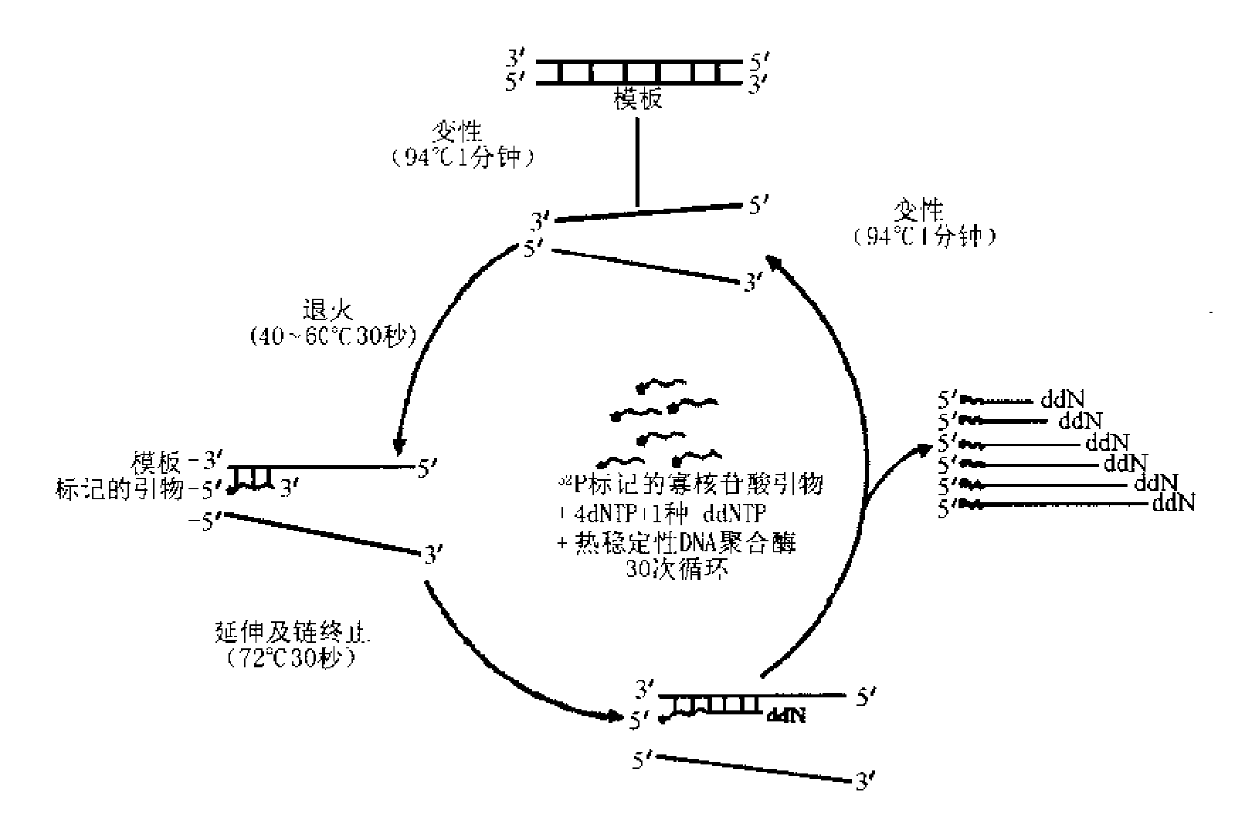


图 8-4 循环测序原理示意图

就可以获得多于 500 碱基的序列。

（2）模板在循环过程中高温变性，不需要变性和乙醇沉淀 dsDNA 模板。双链 DNA 可以直接测序。

（3）高温退火增加了模板与引物退火的严格性。

（4）高变性温度减少了模板 DNA 的二级结构，从而使聚合反应能够越过模板的高级结构区，延伸更完全、彻底，方便了对 PCR 产物、质粒 DNA、λDNA 和粘粒 DNA 等双链模板的序列测定，且更利于对高 G-C 碱基含量模板的测序。

（5）方法自动化程度高，更易操作。

此法适用于 PCR 扩增产物制备的模板、大的 dsDNA 模板如 λDNA、富含 GC 模板及发夹结构丰富的模板的序列分析。根据标记途径，可选用直接掺入标记和末端标记法进行 DNA 序列分析。

8.4.3 测序反应重要试剂

1. 引物

在测序反应中，引物引导新生链的合成。一般测序反应的模板制备是先将目的 DNA 克隆到 M13 或质粒等载体，然后再纯化含目的 DNA 序列的载体 DNA 而获得。因此可以利用目的 DNA 片段两侧翼的载体互补序列设计成测序反应引物，它可以引导任何插入片段的合成，此引物称为“通用引物”或“万用引物”。通用引物一般长为 15~29 个核苷酸，市场已有商业产品，也可以自己合成。

对于以PCR扩增产物为模板的测序，可以用PCR反应的引物作为测序反应引物，也可以是PCR反应产物中互补的一段序列。早期测序试剂盒以引物标记为主，法医学线粒体测序也应用引物标记法，PCR反应引物设计时在引物的3′端连接一M13的通用序列，以便产生测序反应通用引物结合点。目前采用的末端标记法测序，不需要加M13通用引物，因此模板制备的PCR扩增反应引物中不需要连接通用引物，可以自己设计引物序列进行测序，对于新STR基因座作重复序列的测定，可以用STR基因座扩增反应的引物直接进行测序。

引物浓度不宜偏高，浓度过高有两个弊端：一是形成引物二聚体；二是当扩增微量目的DNA并且起始材料又比较粗时，容易产生非特异性扩增产物。一般来说，用低浓度引物，反应特异性较好。

退火温度是PCR测序的最关键因素，高退火温度增加模板与引物杂交的特异性，减少模板的二级结构。引物延伸从退火片段开始，在低温条件下，聚合酶可能遇到强二级结构区。强二级结构可能导致聚合酶的脱落及模板由于聚合酶外切酶活性的裂解，从而在4个泳道上同一位置都出现条带。所以，要用尽可能高的退火温度。长引物和引物高GC含量可以获得较强的信号，最好的结果是应用>24mer引物（GC含量约50%）。

2. 模板

由于PCR灵敏度高，特异性强，它们对模板的要求不是很高，循环测序法能够对几乎未经纯化的微量DNA进行测序。但模板的纯度将影响测序质量，尤其是采用自动测序方案时，这种影响更明显。对污染极为敏感，用通用引物进行PCR测序，微量的载体DNA和克隆DNA的污染将导致测序实验失败。在PCR模板中，不能混有与引物互补的非目的基因序列。传统的测序模板（纯化的M13或质粒DNA）是极好的测序模板，PCR产物、质粒、λDNA也可以进行测序。用本方法测序时，PCR扩增的DNA经柱层析纯化，去除PCR反应后剩余的dNTP和引物。也可以直接取1～2μl未经纯化的PCR产物进行循环测序，但这种测序结果产生的序列梯带的效果差。如果使用低浓度的dNTP（1～20pmol/L）和引物（10 pmol/L）扩增靶DNA，反应中的剩余dNTPs和引物的量减少，直接用PCR产物作为测序模板，可省去纯化步骤，一般可产生高显影度的序列梯带。每次测序需要的模板DNA浓度至少为0.1pmol。更低浓度的模板也可以使用，但其测序结果相对较差。

3. 标记物

高压电泳分离后的DNA片段需要一定的方法来检测结果。最早应用的是放射性同位素标记如$\alpha-{}^{32}P$ - dNTP或$\alpha-{}^{35}S$ - dNTP作为标记示踪物。

放射性同位素有害于健康，因此越来越多地被非同位素标记物地高辛、生物素或荧光标记物所代替。甚至改用银染测序，不需要任何标记物。

4.dNTP类似物

二重对称的DNA区段（特别是GC含量高的）可以形成链内二级结构，在电泳过程中不能充分变性，在电泳过程中不规则迁移，使邻近DNA条带压缩在一起，以致难以读出序列，这种现象不可能通过改变测序反应中DNA聚合酶而得到缓解。但可以通过采用如dITP（2′-脱氧次黄苷-5′-三磷酸）或7-脱氮-dGTP（7-脱氮-2′-脱氧鸟苷-5′-三磷酸）等dNTP类似物进行分辨。这些类似物与普通碱基的配对能力较弱，而且是测序

酶和 Taq DNA 聚合酶等 DNA 聚合酶的合适底物，7－脱氮－dGTP 或 dITP 能缓解一些压缩条带，但并不是所有的均能得到缓解。如果需要采用类似物，首先可试验 dITP，如果用 7－脱氮－dGTP 或 dITP 都无法分辨，可采用反向测序，获得理想结果。

5.DNA 聚合酶

DNA 聚合酶的种类和质量是序列测定成功的关键因素。DNA 聚合酶的种类不同，测序反应的条件可能有所不同。无论何种聚合酶，DNA 聚合酶具有 5′→3′聚合酶活性和 3′→5′的外切酶活性，不具有 5′→3′外切酶活性。DNA 聚合酶有 Klenow 片段、T7 DNA 聚合酶、Taq DNA 聚合酶和测序酶等，由于 Klenow 片段本身的缺陷已不太使用了。

T7 DNA 聚合酶是从 T7 噬菌体感染的大肠杆菌中分离得到的，为一种蛋白复合体，其聚合能力是所有已知 DNA 聚合酶中最强的，且催化合成的 DNA 的平均长度比其他 DNA 聚合酶催化合成的 DNA 要长得多。经过修饰的 T7 DNA 聚合酶不具有 5′→3′外切酶活性，几乎无 3′→5′外切酶活性，是 Sanger 法测序反应中最为理想的 DNA 聚合酶。利用 T7 DNA 聚合酶进行的测序反应可测出 1 000bp 的序列。

Taq DNA 聚合酶具有极好的热稳定性，在 95℃的高温下仍具有稳定的活性，其最佳活性温度为 75～80℃，在测序反应中，它有着其他 DNA 聚合酶所没有的优点——能够测定在 37℃时形成稳定二级结构的单链 DNA 模板的序列；测序结果的带型均一，背景清晰，而且 Taq DNA 聚合酶的持续合成能力较强。

测序酶是经过改造的 T7 噬菌体 DNA 聚合酶，消除了 3′→5′外切酶活性。活性非常稳定，具有很高的链延伸能力和极快的聚合反应速度，是测定较长 DNA 序列的首选酶。测序酶 2.0 版要优于原来的测序酶，测序酶 2.0 版持续合成能力强于测序酶，目前大多数使用的是测序酶。其他的酶有 Stragene 公司的 Pfu 聚合酶，England Biolabs 公司的 Vent 聚合酶，Amersham 公司的 Tub 聚合酶等。

在对 ssDNA PCR 产物测序时，也可以用其他测序酶。

8.4.4 测序反应基本操作步骤

以 Taq DNA 聚合酶循环测序为例阐述基本操作步骤，基本方法也适用于其他的耐热 DNA 聚合酶。但要按所用的聚合酶，相应调整缓冲液的成分及 dNTPs 和 ddNTPs 的比例。产生低本底、高显影度的序列梯带所需的最佳 dNTPs、ddNTPs 的比例受众多因素影响，采用某种方法进行测序时，确定最佳的 dNTP，ddNTP 比例是十分必要的。

（1）引物标记：10～15pmol 测序引物，50μCi［－^{32}P］ATP（比活≥6000Ci/mmol），5μl 10×激酶缓冲液，加水至体积为 50μl，混匀后 37℃预热 5min。加入 1μl 现稀释的 T4 多聚核苷酸激酶（10u），37℃保温 30min，再加入 10u 的 T4 多聚核苷酸激酶，37℃继续保温 30min。Biospin－10 凝胶过滤柱除去游离的［^{32}P］ATP。标记好的引物可放于－70℃保存至少两周。

（2）4 种 dNTP/ddNTP 混合物，每种各取 2μl，分别加入 4 个离心管内，每个离心管中 dNTP 的浓度均为 30μmol/L，每种 dNTP 相应的 ddNTP 浓度分别为：ddATP 1.2mmol/L，ddTTP 1.2mmol/L，ddGTP 90μmol/L，ddCTP 600μmol/L。对富含 G－C 的模板，应用 7－脱氮－2－dGTP 代替 dGTP，以消除电泳时由于压缩现象而产生的假带。

（3）纯化的 PCR 产物 0.1～0.2pmol，5′端^{32}P 标记的测序引物 1～2pmol，3μl 10×测序

反应缓冲液（500mmol/L Tris－HCl，pH8.8，20mmol/L $MgCl_2$），或厂家提供的缓冲液，5u的Taq DNA聚合酶，加水至体积为20μl，混匀。

（4）上述混合物各取4μl，分别加入到步骤（2）准备好的4管dNTP/ddNTP混合物中，混匀后，加20μl矿物油覆盖液面。

（5）将上述4个离心管置于已加热到94℃的热循环仪中，开始循环反应。循环参数：94℃1min，40～60℃30s，72℃30s，循环20～40次。

（6）循环结束后，每管加入4μl终止液（95%甲酰胺，20mmol/L EDTA，0.05%溴酚蓝，0.05%二甲苯青蓝），混匀后，离心数秒。如样品不马上使用，可放－70℃保存一周，但最好在两天内使用。电泳前，将离心管在加热仪上80℃加热3min，取2～3μl加入测序凝胶上样孔内进行电泳。

8.4.5 聚丙烯酰胺凝胶电泳分离片段与读序

反应结束后，将序列反应中产生的各种不同长度的DNA片段按其大小有序地分离开并显示出来。这一步是序列测定成功的关键之一。采用超薄长距离的聚丙烯酰胺凝胶高压电泳进行分离，根据所带标记物的性质进行检测。

1. 变性聚丙烯酰胺凝胶制备

凝胶制备可以在测序反应前一天进行，凝胶灌注后2～24h之间使用均可获得良好效果。凝胶浓度取决于电泳分离的DNA片段大小，一般为6%～8%。当需要测定距离引物50个碱基以内的序列，或化学法测定50bp以下的DNA片段时，应采用高浓度凝胶（10%～20%）。制备步骤同前介绍的PAG制作。

2. 测序产物的凝胶电泳

（1）将凝胶板安装到电泳仪上，鲨鱼齿梳插入凝胶约1mm，倒入1×TBE缓冲液，接通稳压电源，40cm长的凝胶以1 000v进行预电泳20～30min，使凝胶加热到60℃左右。

（2）将测序反应样品在95℃水浴中加热变性2min。关闭电源，再次用注射器冲洗加样孔，用微量移液器上样2.5μl。上样完毕后打开电源开关进行电泳。

（3）电泳时间可根据上样缓冲液中指示剂泳动的位置决定。

（4）取下胶，掀开凝胶上层玻璃板。将凝胶连同下层玻璃板一同放入大瓷盘中，用10%甲醇和10%乙酸溶液固定15min，去除凝胶中的尿素。将一张Whatman MM滤纸或新华3号滤纸（略大于凝胶）铺在凝胶上面，连同玻璃板一同翻转过来，使凝胶落在滤纸上，去掉玻璃板，将一保鲜膜平铺在凝胶上，用手轻轻赶走保鲜膜与凝胶之间的气泡。由于凝胶易破碎，操作时应格外仔细。

（5）将夹在滤纸与保鲜膜之间的凝胶一同放到干胶器上，于80℃真空烤干。去掉保鲜膜，X线胶片直接压在凝胶上，曝光24～36h，得到测序自显影照片。

8.4.6 序列分析

直接阅读序列或将数据输入计算机进行分析。在读取序列时，从胶片的下部往上读取，读出的序列为互补链的5′→3′的序列。

首先从胶片的底端第一条带读起，观察它在A、T、G、C的哪个泳道上或什么颜色，如在G泳道上，则确定此第一个碱基G，然后观察与其紧相邻的一条带，确定是什么碱基，反复如此，阅读所有的条带，得到DNA序列。未知序列至少应读两次，最好由不同

的人来读，然后比较、分析阅读结果。

测序胶片条带的一般规律：

（1）C 的带常常比其他 3 种核苷酸的带弱；

（2）连续出现的多聚 A 中的第一个 A 一般比另几个 A 要强；

（3）连续出现的多聚 C 中的第一个 C 通常比第二个 C 要弱得多；

（4）在 T 之后出现的 G 带较弱。

8.4.7 双脱氧测序法中出现的问题

Sanger 双脱氧链终止法测序操作简单，容易获得结果，但即使是经验很丰富的专家，也不可避免地会遇到各种各样的问题，影响正确结果的获得。在进行序列测定时，除小心操作外，还必须了解各种问题出现的原因，只有这样才能寻求解决对策。

在序列测定的每一步都会遇到各种问题，影响测序结果。在制备好 DNA 模板后，下一步容易出现的问题就是序列胶放射自显影后所显示的序列条带不清楚，例如条带模糊或者同一位置（水平）不止一条带等，这些问题有些不太容易克服，在读胶时允许存在。通过改进制备模板技术或改变链延伸 - 链终止反应中的 ddNTP∶dNTP 比率或聚丙烯酰胺凝胶电泳的条件而加以解决另外的许多问题。对于出现的问题，首先弄清是系统性问题还是模板特异性的问题。下面就测序全过程中容易出现的问题逐一加以讨论。

1. 模板的质量问题

这是较常见的问题。制备模板时应该保证质量。但对模板的质量要求一般由所用的 DNA 聚合酶决定。不同的 DNA 聚合酶对模板要求不同，序列酶对模板的质量要求并不高。用同样的操作程序获得的模板 DNA，用某一种 DNA 聚合酶测序时能得到很好的测序结果，但若用其他酶进行测序反应，几乎无法读取序列。

（1）序列质量很好但条带很弱

① 最可能是模板 DNA 的量较少。延长放射自显影的时间可增加强度，但这样会延迟时间。

② DNA 沉淀中的盐分没清除干净，导致聚合酶的作用受抑制，使谱带发虚，常出现假带。

（2）没有测序模板，结果根本看不见序列带。

（3）假带现象很严重

① 盐分没有除净或污染了聚乙二醇，不仅会产生假带，而且谱带很模糊。增加酶的用量可减少假带数量，但高盐浓度又将降低引物与模板配对的特异性。

② DNA 降解。

2. 引物

（1）序列质量好但条带弱

这种情况通常是引物浓度太低，延伸时温度太低或太高引起的；或引物反复冻融引起降解，从而导致序列带弱。另外，非特异的引物碎片间的退火也使假带的背景增加。引物的贮存液应分装成小包装并保存在 - 20℃，避免反复冻融。

（2）假带多且强

多由于引物浓度过高引起的，假带常常不均匀，不规则。每一批新的引物在使用前都

应该确定最适浓度。

3. 序列反应

(1) 假带很多

①在模板质量保证的前提下，DNA 聚合酶的不当也可引起假带的产生。应保证反应体系是 DNA 聚合酶的最适条件，避免降低酶活性的任何操作。

②如果假带只局限在一条泳道，则常常是由于核苷酸混合液的比例不适合，该泳道带通常比其他的弱。

③核苷酸混合液交叉污染也同样产生假带。

(2) 整块胶的放射性强度分布不均匀

在胶底部的条带强度比胶上部的带高很多，使可读序列的范围缩小。可能是反应体系中核苷酸的摩尔比不平衡引起的，使得多数新生链的延伸提前终止。可比较一系列不同核苷酸浓度比例实验结果，以获到最适比例。严重的比例失调会导致假带的产生。

4. 电泳

电泳过程中，常出现的问题是制胶和加样上的错误。相比之下电泳本身出错机会不多。

(1) 灌胶时形成气泡 玻璃板不洁净或有油脂，灌胶易产生气泡；灌胶时胶液应连续加入，不可暂停或使液流中断。一旦出现气泡，用一细长的 X 射线胶片或更薄的夹条插入到胶液中赶走气泡。

(2) 放射自显影时胶片上出现污点（常伴随细线条带向下延伸） 在倒胶前玻璃板上有尘埃，可用一些质量好的纸巾擦去板上的尘埃。

(3) 带形失真，无法辨认 造成这种现象的原因多种：①凝胶聚合的时间不够。在电泳前应认真检查凝胶是否完全聚合。②配胶的试剂不适当，是较常见的原因，采用电泳级的试剂。在溶解尿素时，不能过度加热，过度加热会降低溶解度，甚至会导致自发的聚合，胶液必须存于 4℃，保存期超过 1 个月就有可能变质。③电泳缓冲液变质可引起条带模糊，几乎整个胶板都受影响。电泳缓冲液应当经常更换。④样品在变性过程中过度处理使条带模糊、强度弱、假带多、泳道的背景高。⑤在加样之前，凝胶中渗出的尿素没有洗净。凝胶中的尿素在很短时间里即可渗出，因此在加样之前动作要快，迅速清洗干净加样孔，并迅速加样、电泳。否则会导致条带模糊甚至凹凸不平。⑥加样量太大也会引起条带模糊，样品在加样孔中的高度应小于加样孔的宽度，否则就过量了。应尽可能迅速地加样并立即电泳以避免 DNA 在加样孔中复性。

(4) 胶走得不齐有弧度 电泳过程中胶板温度过高。采用热循环装置，控制温度。

(5) 凝胶的固定与加压 X 光胶片不当 在固定过程中凝胶中的尿素不能有效地渗出并除尽，会造成条带模糊。在放射自显影时一定要将 X 光胶片与胶紧密地压在一起，否则条带会模糊不清且很弱。

5. 序列本身的问题

(1) 条带强度不均一

① A 带：在连续几个 A 的区域，最下边的 A 带最强。

② C 带：当两个或两个以上 C 带相邻时，自上而下强度逐渐减弱，最下边的 C 带几

乎看不见。C 带本身很弱，尤其 C 带前面是 G 带时更弱。

③ G 带：在 TG 相连的区域中 G 带弱。

(2) 假带 有时也会发生一些序列本身造成的假带，但这不太常见。

① TGCC：这种序列会在 C 泳道或 T 与 G 带之间出现假带。

② GCA：在 A 带的水平线上会出现 T 或 C 的假带。

(3) 压缩现象 这种现象通常出现在 DNA 序列的 GC 丰富区，这些区域往往形成发卡状的二级结构，具有这一结构的片段的电泳迁移率变快，使谱带很集中，压缩在一起，有时甚至造成重叠，难以读出序列。一般情况下，在压缩区带的上方区域谱带的间隙异常变大。解决方法有 4 种：一是电泳时加大电压以提高凝胶的变性温度，减低二级结构的产生，但过高的电压会引起炸板，或者在凝胶中加入甲酰胺使终浓度为 25%～80%，但使用高浓度的甲酰胺会增加背景；二是测定互补链的序列；三是用 ITP 代替 GTP，因为 I:C 配对形成的氢键要比 G:C 配对弱；四是改用 Taq DNA 聚合酶。

8.5 非同位素标记测序及银染测序

自从 Sanger 双脱氧链终止法进行 DNA 测序问世以来，几经改进完善，然而对测序反应产物的检测一直延用放射性同位素的方法，即在链延伸过程中使用放射性同位素标记的核苷酸，使用这种具有危害性的放射性同位素作为标记示踪物，限制了其应用范围。随着非同位素标记及银染技术的发展、完善，非同位素标记测序和银染测序逐渐替代了同位素标记测序。

8.5.1 非同位素标记测序

利用地高辛和生物素测序是近年来发展的一种非同位素测序手段，其酶学部分利用了 Sanger 双脱氧链终止法的测序原理，但在测序反应体系中，以非放射性标记引物（如地高辛、生物素）代替放射性同位素进行测序。将非放射性标记物（如地高辛）与测序引物 5′端共价结合，进行测序反应，反应产物经序列胶电泳分离，通过直接印迹电泳系统（direct blotting electrophoresis system，DBE）将 DNA 条带转移到尼龙膜上，或者通过抽滤转膜的方法将 DNA 条带转移到尼龙膜上，经紫外线照射固定，利用酶标地高辛抗体或亲和素与测序反应产物上的地高辛或生物素结合，最后通过酶显色反应将测序结果显示出来。整个测序过程包括：模板制备；测序反应；电泳及 DBE 转膜；显色反应及序列读取。

8.5.2 银染测序

1. 原理

银染色法是近年来建立的一种非放射性同位素标记的 DNA 序列测定方法。其基本原理是通过高灵敏度的银盐显色步骤来检测末端终止法完成的测序凝胶条带。银染测序系统利用线性扩增，热循环等步骤获得银染检测所需的足够 DNA 量，在一天内完成序列反应，电泳分离，并获得数据。该方法可以在电泳结束后一个半小时内得出结果，同时也省去了放射性同位素操作过程中的特殊防护措施。与荧光染料标记方法比较，荧光及化学发光技术需要昂贵的试剂来标记引物或底物，同时需要昂贵的仪器来检测。银染色测序法使用普通的寡核苷酸引物，无需在引物 5′-OH 末端或底物进行任何修饰，也不需要复杂的仪器设备来检测结果，可以在显色后的胶上直接读序。该法快速、安全、费用低，而成为酶法

测序分析中的重要测序系统。银染测序包括：模板制备；测序反应；序列胶电泳分离；银染显色结果与读序。银染法测序凝胶制备方法同 STR 复合扩增产物银染检测，采用特殊方法分别处理测序凝胶的长短两块玻璃板，使电泳结束后凝胶能够附着在短玻璃板上，而与长玻璃板能够比较容易分离，凝胶在短玻璃板的托载下，能够方便地进行显色操作。

8.6 DNA 自动测序技术

DNA 测序技术经过十几年的发展，不断改进和完善，建立了 DNA 自动测序技术，使操作更简便，省时，结果更为清析和易于读解。目前已为世界上许多实验室所采用。下面主要介绍 ABI 自动 DNA 测序仪的基本原理。

8.6.1 DNA 自动测序技术原理

1. 化学原理

ABI 公司研究人员经过长期实验摸索，最终确定了 4 种荧光发色基团分别作为 DNA 聚合酶链反应中 4 种双脱氧（ddNTP）终止的 DNA 片段的标记物，即 4 种 DNA 片段带上了不同的颜色。当将这些 DNA 片段混合在一起加入同一样品孔中，经电泳展开，相互间仅差 1 个碱基的 DNA 片段，便形成一条连续的且具有 4 种颜色的阶梯分布图象。每一阶梯上的 DNA 片段，其 3′端终止碱基的特征由标记在该 DNA 片段上的特征荧光发色基团发出的荧光波长所指示。

（1）荧光发色基团的选择标准

自动测序利用不同荧光染料标记来检测片段，对于荧光染料的选择标准同在 STR 章节介绍的情况。

（2）荧光基团掺入的方式：

荧光基团掺入的方式：①将荧光素预先标记在测序反应所用引物的 5′端。相同碱基排列的寡核苷酸片段作为骨架被 4 种荧光素标记后，便形成了一组（4 种）标记引物，这种方式被称为引物标记（Dye - Primers）。在测序反应中，与传统手工测序操作一样，模板、反应底物、DNA 聚合酶及引物等按 A、C、G、T 编号被置于 4 支微量离心管中。特定荧光标记的引物则与特定的双脱氧核苷酸底物保持对应关系。以 A 管为例，商品名为 JOE 的引物（pUC/M13 通用测序引物 5′TGTAAAACGACGGCCAGT3′）是与 ddATP 加在一起的。这样以 ddATP 终止的所有 DNA 片段 5′端都标记了相同的荧光基团（图 8 - 5）。②将荧光素连在终止底物的双脱氧单核苷酸上。在反应中带荧光发色基团的 ddNTP 在掺入 DNA 片段导致链延伸终止的同时，使该片段 3′端又标上一种特定的荧光素。4 种荧光素分别与 4 种 ddNTP 底物连接，反应结束时产生的四组 DNA 片段分别由特定 ddNTP 所终止，并且标记有 4 种不同的荧光发色基团。这种方式被称为末端标记（Dye - Terminators）（图 8 - 6）。

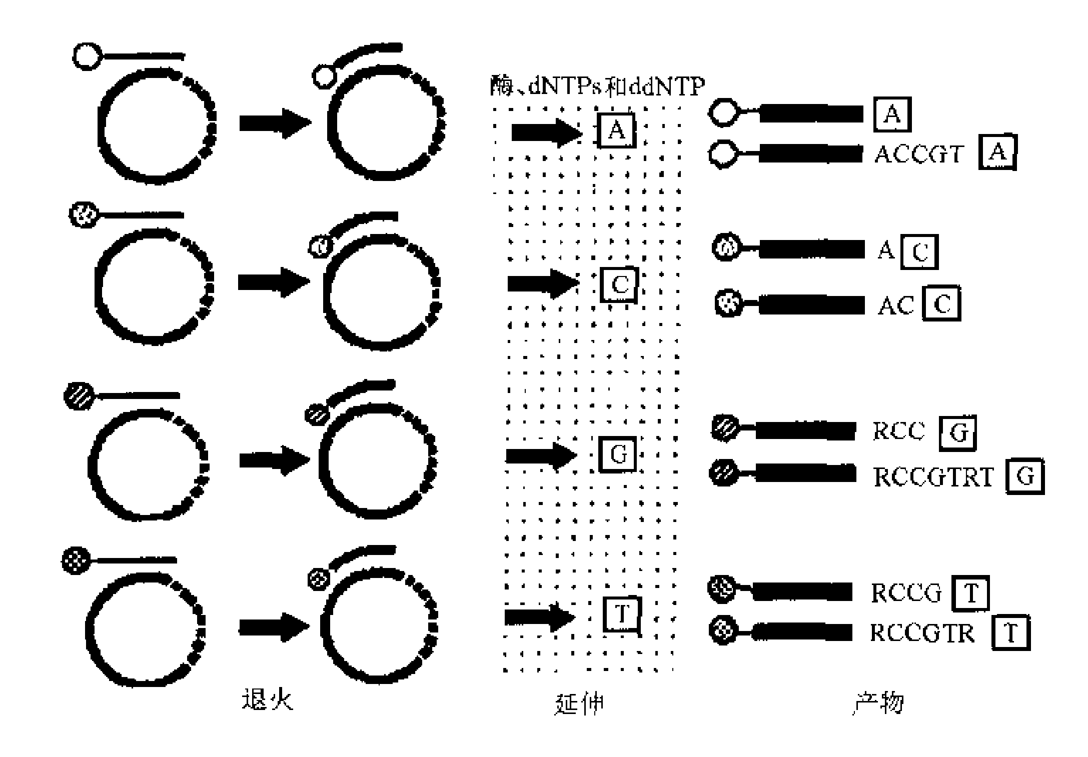


图 8-5　荧光标记引物化学原理

○表示5′端带不同荧光染料的引物，A等表示终止底物ddNTP

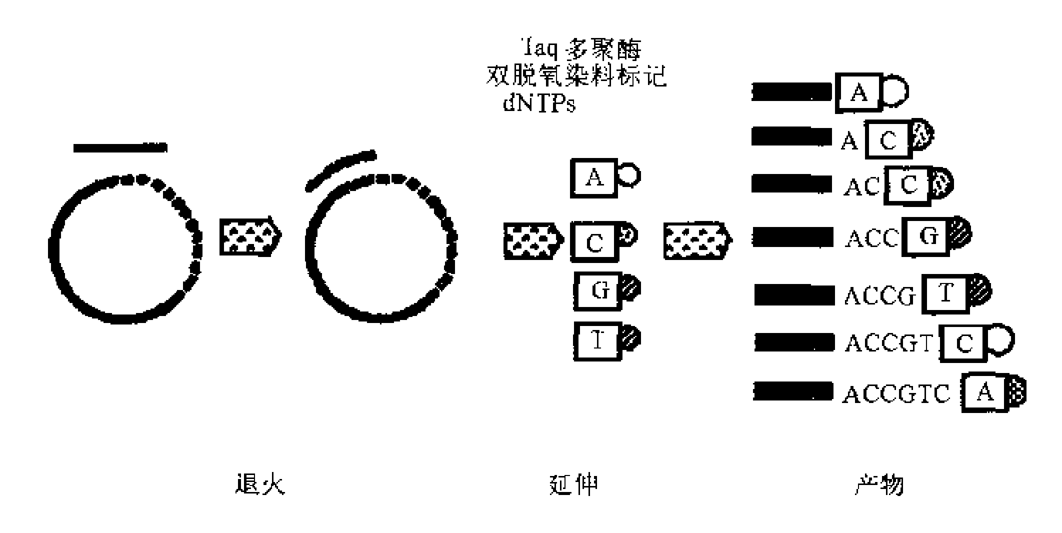


图 8-6　荧光标记终止底物法化学原理

A○等表示带不同荧光染料的终止底物（ddNTP）

以上两种策略都确定了4种荧光素与4种双脱氧单核苷酸所终止的DNA片段之间的一一对应关系，这是随后电泳中信号检测以及最终阅读序列的基础。两种掺入方式的区别在于：引物染料标记使荧光发色基团标记在长短不同的DNA片段的5′端，也就是荧光素标记过程与延伸反应终止分别发生在同一DNA片段的另一端，且标记发生在引物与模板

的退火过程；而末端染料标记法使标记与终止过程合二为一，采用染料末端标记法，A，C，G，T 4 个测序反应可以在同一个反应管中完成，减少了加样的步骤，那种由于酶过早从模板中脱落下来引起的延伸假停止也检测不到了，因为这些片段没有染料标记。采用引物染料标记法，则要求 A，C，G，T 4 个测序反应分开进行，但这种方法得到的每个碱基测序信号强度比染料末端标记法更均匀，而且一般序列阅读较长。表 8－1 为两种标记方法及聚合酶的一些比较。

表 8－1 各种测序试剂的特征

	染料引物标记循环测序	染料末端标记循环测序	染料引物标记测序酶	染料末端标记测序酶
正确性（24cm）	450 碱基，98%	450 碱基，98%	450 ～ 500 碱基，98%	ss：450 ～ 500 碱基，98% ss：350 ～450 碱基，98%
主要优点	信号平衡，模板量少	单管反应。模板量少，没有假停止，背景低	信号非常平衡	单管反应，信号相对较平衡，没有假停止
注意点	偶尔假停止和丢带，有时看见带压缩	必须去除过量的终止物	一般信号较弱，可能需要更多的模板，可能有假停止，只能用单链模板	双链 DNA 需要变性，dsDNA 只能测到 350 个碱基，GC 含量高的模板出现带压缩现象
每个测序的反应管数	4 个	1 个	4 个	1 个
模板	20 ~ 200ng	20 ~ 200ng	用磁珠纯化	0.3 ~ 10μg
PCR 产物				
ss	0.6μg	0.25 ~ 0.50μg	2.0μg	2.0μg
ds	1.2 ~ 1.5μg	1.0μg	不能用	5.0μg
信号的平衡性	平衡	平衡性差	非常好	平衡
修饰染料	FAM，JOE，TAMRA，ROX（fluorescein 和 rhodamine）	R110，R6G，TAMRA，ROX（rhodamine）	FAM，JOE，TAMRA，ROX（fluorescein 和 rhodamine）	FAM，HEX，NAN，JOE（fluorescein）
标记位置	引物的 5′端	ddNTP	引物的 5′端	ddNTP
酶	AmpliTaq[R]	AmpliTaq[R]	Sequenase[R]	Sequenase[R]

2. 仪器原理

ABI 公司已推出的两种型号（373A 和 377A）自动 DNA 序列仪的主机是一台带有自动检测系统的高压电泳装置。在此系统中，不同 ddNTP 终止的 DNA 片段由于标记了不同的荧光发色基团，它们可以混合起来加样在同一样品孔中。两极间极高的电势差推动着 DNA 片段在凝胶中泳动并达到相互分离，并依次穿过设置在靠近胶板下端的固定检测区域。图 8－7 是窗口处激光束光学路线分布以及荧光信号收集检测系统的示意图。由激光器发出的极细光束，通过精密的光学系统被导向凝胶表面的检测区，在这里激光束以与胶板垂直的角度和穿过“窗口”的 DNA 片段相遇；DNA 片段上的荧光发色基团吸收了激光

束提供的能量而发出特征波长的荧光。这种荧光又经过聚合后传给一个高灵敏度光电管而转化为电信号，4 种不同波长的荧光由安放在光电管前方的旋转滤光片加以区分。DNA 片段电泳与检测是一个动态过程，DNA 片段通过窗口时信号被自动记录，继续向胶板底部泳动，最后进入下层缓冲液中，这就是自动测序系统中广泛采用的真实时间扫描技术(real time scanning)。实验者所感兴趣的是那些 DNA 片段在到达距加样孔 25cm 处的“窗口”检测区时，大小仅相差一个碱基的 DNA 片段间有足够的分离。一个 DNA 片段经过“窗口”时，它的特征荧光波长和强度被及时记录下来，随后进入窗口位置的是长一个碱基的片段。这就是说，DNA 序列的读解是通过记录经过窗口区域的 DNA 片段的特征波长及顺序来实现的，测序与电泳步骤同步完成。

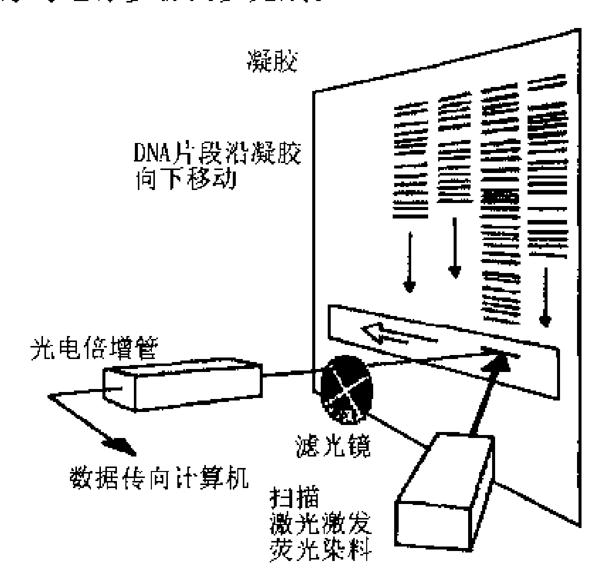


图 8-7 荧光 DNA 分析原理图

单道电泳测序（one lane sequencing）是 ABI 公司自动测序装置上的又一改进。在放射性自显影测序工作中，碱基顺序的确定需要实验人员细心地比较 4 条电泳道上显色条带的分布。有时因温度梯度或凝胶不均匀使胶板上不同区域 DNA 片段泳动不均匀，都将给解读工作带来很大困难。采用单道电泳技术，所有的 DNA 片段由一条电泳道依次达到检测窗口，便消除了上述因素造成的影响。即使因边缘效应等问题造成 DNA 条带倾斜，也不会影响信号的判读。

3. 数据收集和分析

由主机收集的荧光信号在转化为电信号后被输送给计算机储存和进一步处理。当电泳-检测过程终止时，计算机便建立了一个称为“GEL FILE”的总文件。它实际上是一个极为复杂的多维矩阵数据库，包含了各荧光信号的波长特征及强度、空间座标等参量。计算机软件以此为基础，在屏幕上显示出凝胶板上每个样品经电泳分离后 DNA 片段排列情况的模拟图象，在其中用 4 种颜色条带排列表现各 DNA 片段终止碱基的种类和排列次序。

8.6.2 测序方法

用 PCR 方法扩增出测序模板，纯化模板。准备测序胶，进行测序反应，最后上机测序。其具体方法将在线粒体 DNA 测序章节中作详细介绍。

8.6.3 核苷酸顺序的阅读

自动测序连续采集的数据在计算机中储存并进行迅速处理，结果清晰准确，分辨率高，分析人员根据显示的四色图象准确分辨出 DNA 片段的碱基排列顺序。根据计算机软件中碱基的代表符号，结合扫描图，进行阅读和编辑。大多数软件采用国际生化联合委员会标准碱基符号（IUB Standand nucleotide Codes）（表 8－2）。

表 8－2 国际生化联合委员会标准碱基符号

符号	碱基	含义
A	A	腺嘌呤
C	C	胞嘧啶
G	G	鸟嘌呤
T 或 U	T 或 U	胸腺嘧啶
R	A 或 G	嘌呤
Y	C 或 T	嘧啶
S	G 或 C	强结合
W	A 或 T	弱结合
K	G 或 T	含酮基
M	A 或 C	含氨基
B	C 或 G 或 T	非 A
D	A 或 G 或 T	非 C
H	A 或 C 或 T	非 G
V	A 或 C 或 G	非 T、U
N 或 X	任意碱基及空位	任意或未知

8.7 化学降解测序法

化学降解测序技术是由 Maxam 和 Gilbert 在 1977 年创立的。其主体思想与 Sanger 链终止测序法相同。两者都产生以 A 为一个系列、G 为一个系列、C 为一个系列和 T 为一个系列的不同长度的 DNA 片段。区别在于前者的新 DNA 片段以合成的方式产生，而后者的新 DNA 片段不是合成产生而是降解原有的 DNA 链。

以 G 系列为例阐明该法的原理。首先将准备测序的 DNA 链用同位素进行末端标记，用对 G 有特异修饰的试剂——二甲基硫进行碱基修饰，在这一步 DNA 链仍保持完整。第二步加人第二种化学试剂哌啶，在与 G 核苷酸（修饰过的）相连接的磷酸基团处将 DNA 链切断产生不同长度的片段。每一个裂解产物都有两个片段，其中一个片段带有标记物，可在放射自显影中显现，而另一片段由于没有标记物将不能被观察到（图 8－8）。以同样的方式产生 A、C、和 T 的系列片段，比较它们的序列，得到待测片段序列。化学降解测序技术目前较少应用，本章不作详细介绍。

8.8 固相微型测序

固相微型测序是 Syvanen 等 1990 首先报道应用的，用来分析 ApoE 基因。顾名思义，这种测序方法结合一个固相化过程，进行部分测序。固相微型测序法可用于法医学个体间的序列比较、医学疾病诊断、只有单个或数个碱基差异的序列分析或已知多态性碱基位置

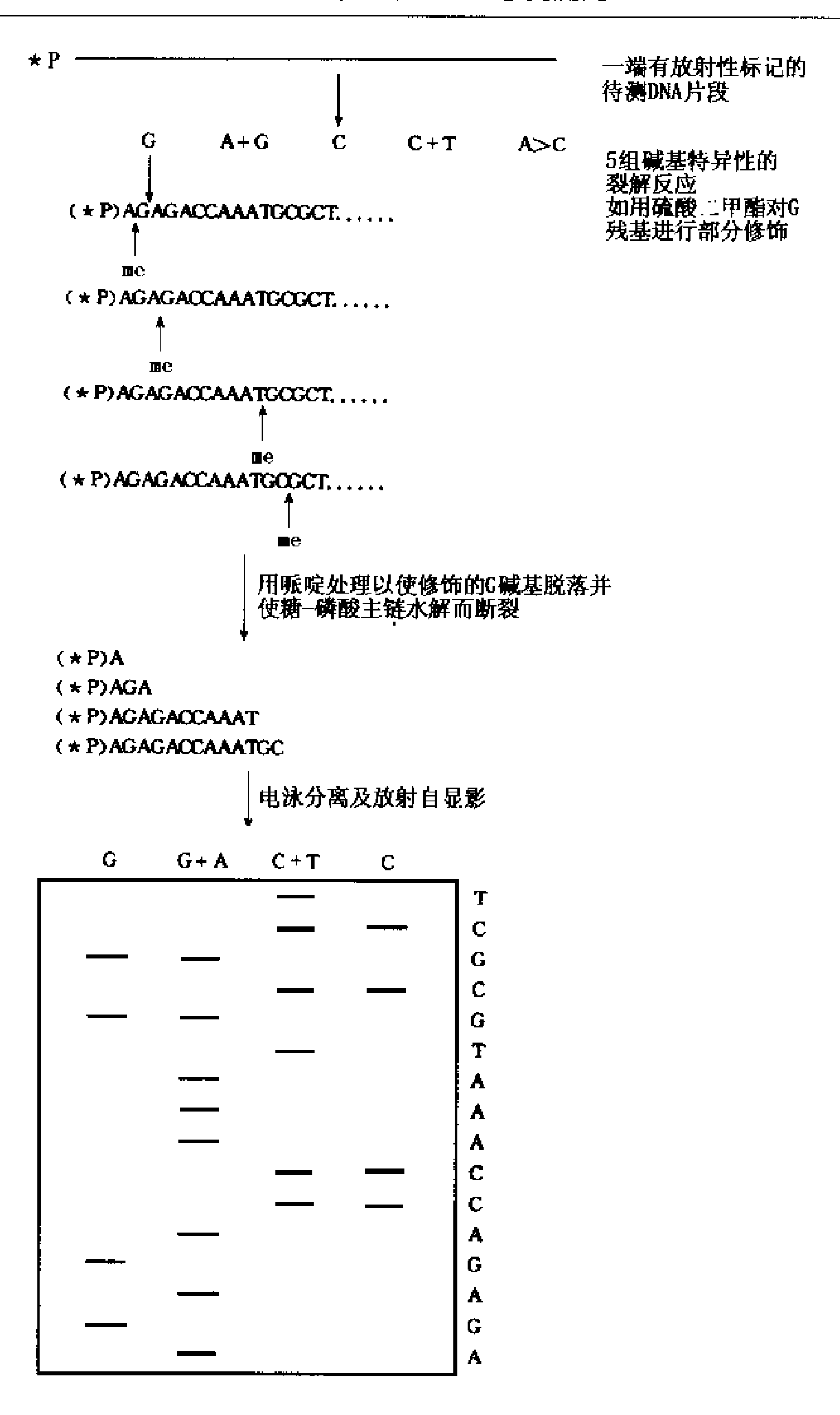


图 8-8　M&A 化学法测序原理示意图

的分析等。

固相微型测序的原理如图 8-9 所示。

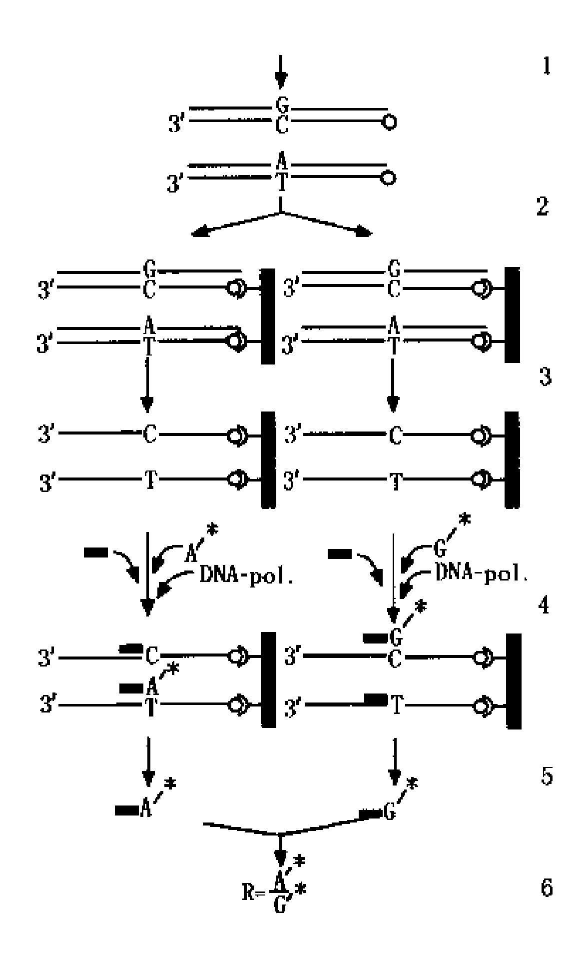


图 8-9 固相微型测序的原理

(1) 生物素标记引物和非生物素标记的引物扩增 (2) 生物素标记产物亲和结合到链亲和素包埋的微孔滴定板 (3) 洗涤并用碱变性 (4) 微型测序引物延伸反应 (5) 延伸产物测定 (6) 结果解释

(1) 首先用一条 5′端带生物素标记的引物和另一条不带生物素的引物扩增出含多态性位置的序列; (2) 带生物素的扩增片段通过亲和素或链亲和素包埋的固相支持物被固定捕捉; (3) 通过洗涤除去多余的 PCR 试剂, 固定的 DNA 模板通过碱处理变成单链得到单链测序模板; (4) 单链模板以多态性碱基紧邻的互补序列为引物进行微型测序反应——链延伸-终止一步反应; (5) 再通过碱处理将结合的带标记的引物延伸链从固定模板中洗脱出来, 液闪计数或分光光度法测定、或酶化学发光检测、或凝胶电泳结合荧光检测引物链, 确定单链 DNA 片段中的碱基序列。

微型测序反应利用的也是 DNA 聚合酶准确的链反应与链终止, 由于底物的缺乏或添加双脱氧核苷酸底物导致链延伸的终止。不同的是采用的引物序列设计特殊, 引物的 3′最后一碱基恰好是变异碱基, 延伸反应添加的第一个碱基即为多态性碱基, 反应底物为一种脱氧核苷酸或双脱氧核苷酸, 引物退火后, 在 DNA 聚合酶催化下加入该种碱基进行链的延伸, 由于底物中只有多态性碱基中的一种碱基, 链的延伸就会被迅速终止。如果用带标记的双脱氧核苷酸, 一个碱基添加后链的延伸马上终止, 得到长度为引物长加 1 个碱基的产物链。早期的同位素方法使用的是脱氧核苷酸, 后来发展的荧光标记检测的则使用四色

荧光标记的双脱氧核苷酸，一个样品只需要一个测序反应就可以迅速得到结果，大大提高工作效率。而且采用多色荧光标记，可以复合检测多个核苷酸的多态性，如线粒体 DNA 中 12 个多态性核苷酸。固相化微型测序操作将在线粒体部分介绍。

固相微型测序进行染色体 DNA 二等位基因（即单核苷酸多态性）遗传标记的分析方法已成功地应用到法医学中。通过电泳分离四色荧光，可以同时检测 12 个 DNA 遗传标记，这 12 个遗传标记在芬兰人群中个体识别率为 0.99996，平均非父排除率为 0.90，一次进行多个分析可以减少成本、节省人力及时间，该法从 PCR 一直到加样电泳前均在一个反应管中，很适合自动化分析。

第九章 DNA 芯片技术

DNA 芯片的概念来自于计算机芯片。所谓 DNA 芯片（DNA chip 或 microarray（微阵列））是指利用大规模集成电路的手段控制固相合成成千上万个已知序列的 DNA 或寡核苷酸探针，并把它们有规律地排列在面积不大的的硅片、玻璃片、聚丙烯或尼龙膜等固相支持物上，然后用荧光染料等标记将要研究检测的材料如 DNA、RNA 等，与芯片上的探针杂交，再通过激光共聚焦或 CCD 扫描仪对芯片 DNA 的荧光强度进行扫描，并配合计算机系统对每一个探针上的荧光信号作出比较和检测，从而迅速得出所需要的信息。DNA 芯片，它是生物芯片中的一种，也是最早开发的生物芯片。芯片技术用于检测的时间只需 30min，甚至更少，比常规方法快几十至几千倍。基因芯片主要利用芯片技术中信息的集约化和平等处理，具有无可比拟的高效、快速和多参量的特点，是传统生物技术如检测、杂交、分型和 DNA 测序的一次重大创新和突破。DNA 芯片发展至今不过 10 年左右，但进展神速，迄今已有近百家公司从事生物芯片相关工艺、设备及检测手段和软件的开发。

9.1 DNA 芯片的分类与制作

9.1.1 DNA 芯片的分类

根据芯片或微阵列上排列的探针种类的不同，可将 DNA 芯片分为两类：

(1) 寡核苷酸探针　寡核苷酸原位合成或合成后固定在芯片上，根据杂交信号出现部位的寡核苷酸序列推测与其互补的待测序列。寡核苷酸芯片可用于突变检测、表达监控、基因发现和遗传制图等。

(2) cDNA 芯片　将 cDNA 固定在玻片或芯片等固体表面。cDNA 可来自于感兴趣的 cDNA 克隆、cDNA 文库或开读框。cDNA 芯片可用于大范围筛查和基因表达的研究。有时也可应用基因组 DNA 或经基因组错配扫描纯化的 DNA 替代 cDNA 制造芯片。

根据杂交机理，分为主动性的电杂交芯片和被动杂交芯片，前者有如 Nanogen 公司生产的芯片，探针主动富集待测 DNA 分子，加快杂交速度，在数分钟内可以完成杂交过程。大多数杂交芯片则为后一种芯片，利用载玻片、尼龙膜等载体的芯片。

9.1.2 DNA 芯片的载体材料

1. 载体的要求

制作 DNA 芯片的载体要求是：①载体表面必须具有可以进行化学反应的活性基团，以便与生物分子进行偶联；②使单位载体上结合的生物分子达到最佳容量；③载体应当惰性、有足够的稳定性，包括机械的、物理的和化学稳定。惰性，指的是载体其它性能或特异性吸附度不应该干扰生物分子的功能。稳定性，是指在进行分子杂交或结合时，对可能遭受的一定压力或酸、碱条件而不发生变化；④载体具有良好的生物兼容性。

2. 载体材料

制作 DNA 芯片的载体材料很多，大致可分为 4 类：无机材料、天然有机聚合物、人工

合成的有机高分子聚合物以及各种高分子聚合物制成的各种膜。总体说来，虽有几十种甚至上百种载体材料之多，但目前只有少数几种，像玻璃片、金属片和各种有机高分子制作的薄膜等适用于制作 DNA 芯片。其中玻璃片受到各国研究者的重视，其原因：一是来源方便，表面羟基可以用 N，N－二乙氧基氨氧基三乙氧基硅烷作表面处理，然后能够偶联核酸，也可以用巯基标记的生物分子直接和玻璃表面作用等。二是大多数生物芯片采用了发光检测的方法，不管透射光还是反射光它均适合。三是可以用光刻的方法在玻璃芯片上刻出微流路，或者它与刻好微流路的硅橡胶压紧联合使用的微流路，通过微量注射泵在载体上进行生物分子的联接、DNA 杂交、清洗和检测等。表 9－1 为几种载体的性能比较。

表 9－1 芯片载体

片基类型	片基举例	探针固定方式	探针密度	显色及检测方式
刚性片基	玻片、半导体硅片等	原位合成 也可预先合成后点样	高	荧光激光共聚焦扫描、生物传感器等
薄膜片基	硝酸纤维膜、尼龙膜等	预先合成后点样	低	化学显色或化学发光

3. 载体的活化

活化，是指载体的修饰过程，即通过化学反应用活化试剂在载体表面键合上活性基团，以便与配基共价结合，形成具有生物特异性的亲和载体，用来固定 DNA。

活化的方法很多，但是到目前为止还没有一种活化方法是通用的、适合所有载体的活化和配基的偶联。不同载体用不同的活化试剂进行载体表面活化，例如表面为羧基的载体，用 EDC 和 NHS 把表面变成活泼酯，然后再同各种不同的生物分子偶联。采用类似的方法，可以分别把羟基、氨基、醛基、肼基、巯基等的载体表面活化成活性基团，然后再进行不同的键合反应（图 9－1）。

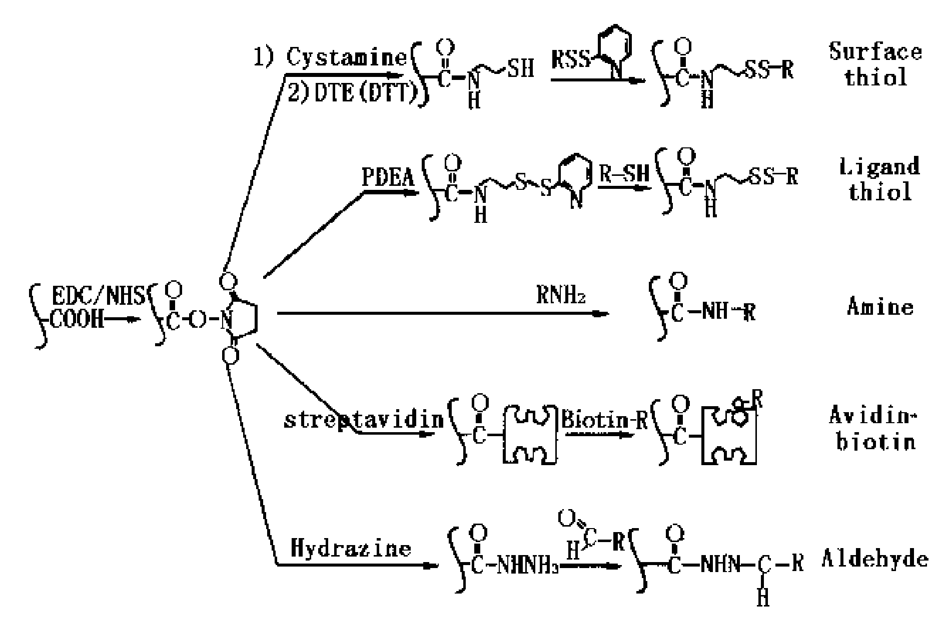
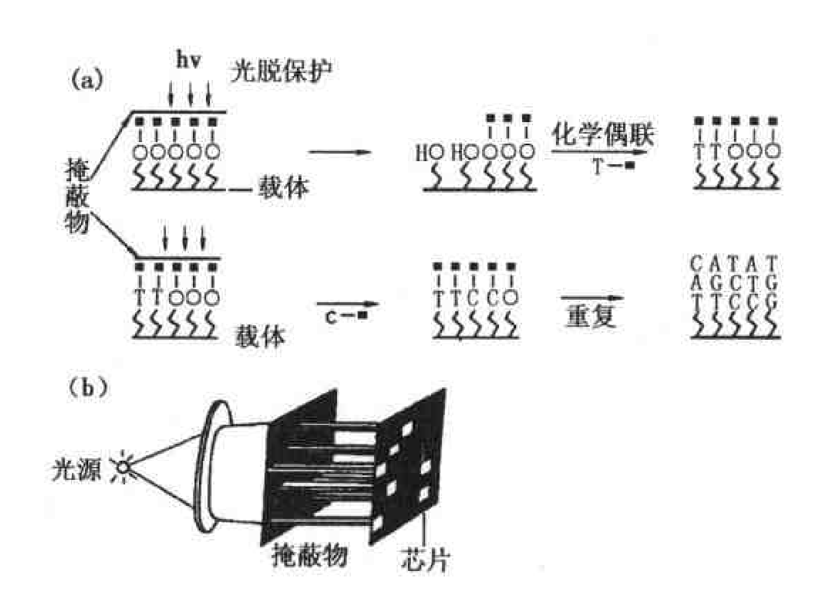


图 9－1 载体表面的活化

9.1.3 DNA 芯片制备的基本途径



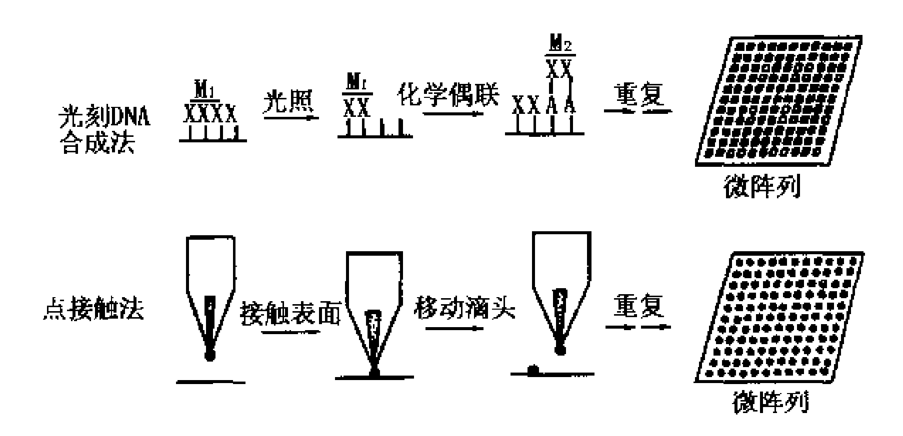


图 9-3 二种制作芯片的方法

列，如图 9-4。

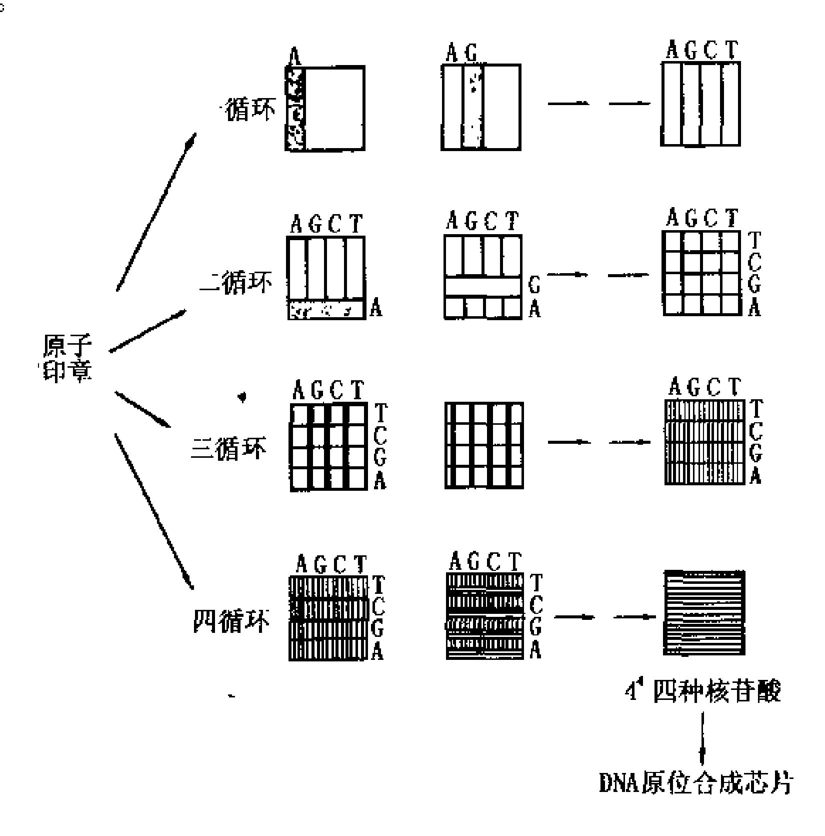


图 9-4 原子印章法 DNA 芯片

4. 直接点样

直接点样法是 Shalon 和 Brown 等人建立的。将合成好的探针、cDNA 或基因组 DNA 通过特定的高速点样机器（图 9-5）直接点在玻片等固体上。优点是实验样机很容易做，可以在 3.6cm^2 面积内点上 10 000 个 cDNA。不足点是每个样品都必须预先合成和纯化好。

5. 电子芯片

它是由美国 Nangen 公司开发的。这种芯片为带有正电荷的硅芯片，芯片经热氧化，

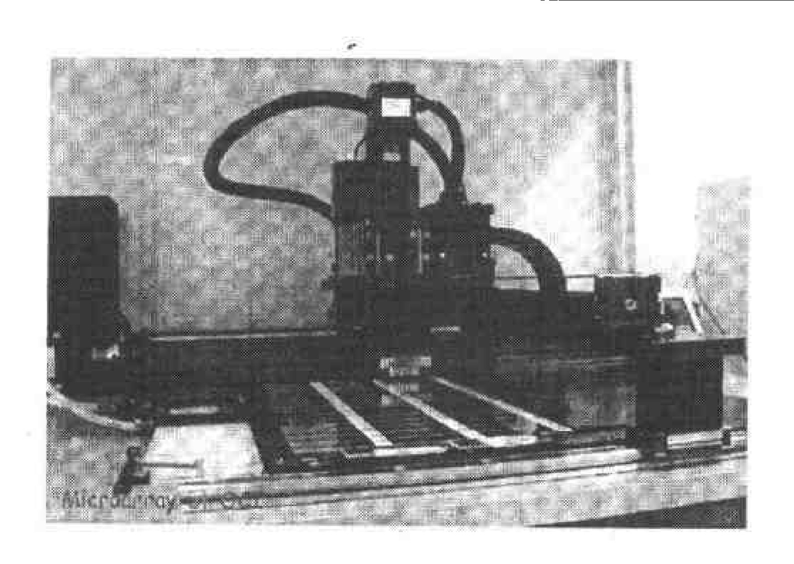


图 9－5 点样仪

制成 1mm 的阵列，每个阵列含有多个微电极，在每个电极上通过氮化硅沉积和蚀刻制备样品池。将连接亲和素的琼脂糖覆盖在电极上，在电场作用下生物素标记的探针即可结合在特定电极上。电子芯片的最大优点是杂交速度快，可在几分钟内完成，可大大缩短分析时间。不足点是制备复杂、成本高。

6. 三维芯片

它是由美国的 Packard、Motorola、Argoone 三家国家实验室机构与俄罗斯 Engelhardt 分子生物学研究所合作开发的一种芯片技术。在一片显微镜载玻片上，有 10 000 个微小聚丙烯酰胺凝胶条，每个凝胶条可用于靶 DNA 的分析。先把已知化合物加在凝胶条上，用 3cm 长的微型玻璃毛细管将待测样品加到凝胶条上。每个毛细管能把小到 0.2nl 的体积打到凝胶上。三维芯片具有其它生物芯片系统不具有的三个优点：一是凝胶条的三维化能加进更多的已知物质，提高了敏感性。二是可以在芯片上同时进行扩增与检测。一般情况下，必须在微量多孔板上先进行 PCR 扩增，再把样品加到芯片上，因此需要进行许多额外操作。而且三维芯片所用凝胶体积很小，使 PCR 扩增体系的体积减小 1 000 倍（总体积约 nl 级），从而节约了每个反应所用的 PCR 酶（约减少 100 倍）。三是三维构象形式存在的基因材料以其天然状态在凝胶条上进行杂交分析。

7. 流过式芯片

流过式芯片（flow－thru chip）是 Gene Logic 开发的一种芯片，在芯片载体上制成格栅状通道，设计及合成特定的寡核苷酸探针，结合于微通道内的芯片的特定区域。从待测样品中分离 DNA 并对其进行荧光标记，然后，该样品流过芯片，固定的寡核苷酸探针芯片捕获与之互补的 DNA，最后用 Gene Logic 的信号检测系统分析结果。流过式芯片可用于高通量分析已知基因的变化，具有以下三个特点：（1）敏感性高。由于寡核苷酸吸附表面的增大，流过式芯片可监测稀有基因表达的变化。（2）速度快。微通道加速了杂交反应，减少了每次检测所需时间。（3）价格降低。由于采用了特殊的共价化学技术将寡核苷酸吸附于微通道内，使每一种流过式芯片可反复使用多次，从而使其成本降低。

9.2 DNA 芯片检验的基本原理与程序

1.DNA 芯片识别的基本原理

DNA 芯片检验可以说是固相反相杂交，是无数个固相杂交的组合，借助于计算机进行分析。待测 DNA 经分离、扩增及标记荧光物质后，在一定条件下与 DNA 芯片上的分子按碱基互补配对原理杂交结合，将杂交后的芯片放进扫描仪扫描。带有荧光标记的 DNA，受光束激发后发出荧光，荧光的强弱与 DNA - 探针杂交的程度有关。通过计算机处理即可给出目的基因的结构或表达信息。

2.DNA 芯片的检验程序

主要包括：(1) 根据检验或研究需要制备和选择 DNA 芯片；(2) 将目的 DNA 用荧光标记，制成靶 DNA；(3) 杂交：即将靶 DNA 加在 DNA 芯片上与芯片上 DNA 杂交；(4) 荧光检测：即应用激光共聚焦或 CCD 扫描仪扫描杂交结果；(5) 计算机处理阅读结果。

杂交条件的选择与试验目的有关，多态性分析或者基因测序时，每个核苷酸或突变点都必须检测出来，通常设计一套四种寡聚核酸，在靶序列上跨越每个位点，只在中央位点碱基有所不同，根据每套探针在某一特定位点的杂交特异性，即可测定出该碱基的种类。如果芯片仅用于突变检测，要鉴别出单碱基错配，需要更高的杂交特异性和更短的时间。

此外，杂交反应还必须考虑杂交反应体系中盐浓度、探针 GC 含量和所带电荷、探针与芯片之间连接臂的长度及种类、检测基因的二级结构的影响。有资料显示探针和芯片之间适当长度的连接臂可使杂交效率提高 150 倍，连接臂上任何正或负电荷都将减少杂交效率。由于探针和检测基因均带负电荷，因此影响他们之间的杂交结合，为此有人提出用不带电荷的肽核酸（PNA）做探针。虽然 PNA 的制备比较复杂，但与 DNA 探针比较有许多特点，如不需要盐离子，因此可防止检测基因二级结构的形成及自身复性。由于 PNA - DNA 结合更加稳定和特异，因此更有利于单碱基错配基因的检测。

9.3 杂交芯片信号检测

9.3.1 扫描仪的分类

DNA 芯片与带标记的目的 DNA 杂交后，必须用显色或扫描测定信号。以尼龙膜为载体的芯片，常用化学显色或发光反应检测结果。对于其他类型的 DNA 芯片，大都采用荧光标记，需要荧光扫描仪扫描检测信号。

根据芯片扫描仪采用的光电耦合器件可以将芯片扫描仪分为光电倍增管型和 CCD 型。根据激发光源的不同还可将芯片扫描仪分为激光型和非激光型。DNA 芯片多选用激发波长较长、斯托克斯位移较大的荧光染料标记，以提高芯片测定的稳定性、降低背景、提高灵敏度。激发光的波长一般为可见光区，多用绿色 532nm 和红色 635nm 激光器。

目前商业化芯片扫描仪主要有两大类型：激光共聚焦芯片扫描仪和 CCD 芯片扫描仪。

9.3.2 芯片扫描仪原理与特性

1. 激光共聚焦芯片扫描仪

激光共聚焦芯片扫描仪以激光作激发光源，发射的荧光强度高，检测灵敏度高。由于 DNA 芯片分析中，靶分子常常用 2 种或 2 种以上的荧光染料标记，芯片扫描仪多采用 2 种

或2种以上不同波长的激光器作激发光源，最多达4种，激发不同荧光染料标记的靶分子。常用的激光器有氩离子激光器、氩氪离子激光器、氦氖激光器和固态激光器。波长可以从488nm至近红外区。

芯片扫描仪的设计原理类似于激光共聚焦显微镜，其基本原理如图9-6。由激光器发射出来的激光由透镜组先扩展成直径较粗的光柱，经激发窄带干涉滤光片（带宽通常为几十纳米），过滤除去其他波长的光，这样可大大降低检测的背景，再由透镜组重新聚焦后，由二色镜反射至物镜组，二色镜设计成这样的功能：入射激光可以被反射，发射的荧光可以通过，但不被反射。物镜组将激光聚焦在芯片上，聚焦后的光斑一般从几个微米到几十个微米，标记有荧光染料的靶分子在激光激发下产生的荧光由物镜组捕获后变成平行光，通过二色镜，由反射镜（也可设计成不用反射镜）反射至发射窄带干涉滤光片（带宽通常为几十纳米）以滤除发射荧光以外的光，再由透镜组聚焦在共聚焦光阑上，通过光阑的光由光电倍增管PMT检测经放大器放大，放大信号由A/D转换器将模拟信号转换成数字信号后，由计算机采集存贮。扫描的结果由计算机采集并以TIFF（Tagged Image File Format)）或GIF（CompuServe Graphics Interchange Format）或JPEG（Joint Photographic Experts-Group Format）等图像文件格式以16位文件存贮。采集得到的各种荧光染料标记的芯片图像文件，可以由软件在计算机中进行合并处理。这类芯片扫描仪典型代表为ScanArray™系列产品，3000，4000，5000型。

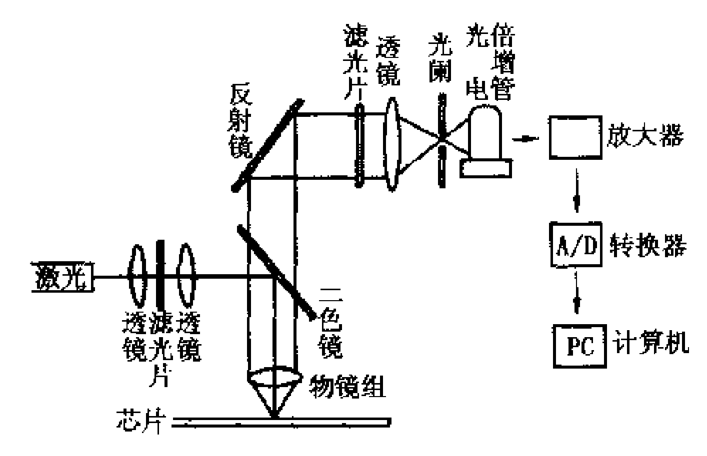


图9-6 激光共聚焦芯片扫描仪原理图

激光共聚焦芯片扫描仪由于采用激光作激发光源，光电倍增管检测，因此具有较高的灵敏度，可检测每平方微米零点几个荧光分子。激光共聚焦芯片扫描仪完成1次扫描一般需要5min左右。

采用激光作激发光源的芯片扫描仪除了激光共聚焦芯片扫描仪外，还有激光激发的扫描仪，采用双固态激光器532nm、635nm作激发光源，采用一种Flying Objective专利技术采集荧光，可提高采集效率，荧光用光电倍增管检测。

2.CCD芯片扫描仪

CCD芯片扫描仪相对于激光共聚焦芯片扫描仪来说，结构比较简单，其激发光源多采用氙灯或高压汞灯，单色系统采用激发窄带干涉滤光片，一般有数片，由电机驱动选择。

CCD 为 CCD 摄像头，镜头前按有电机驱动选择的发射窄带干涉滤光片。其基本原理可参见图 9－7。DNA 芯片在经激发窄带干涉滤光片滤光后的单色光激发下产生荧光，经发射窄带干涉滤光片由摄像镜头捕获成像在 CCD 上。图像信号可由 CCD 摄像头直接传送到图像卡，图像卡一般为 PCI 卡，插在计算机的 PCI 插槽中，图像卡将信号转变成数字信号。由计算机存贮，存贮的文件格式为 16 位 TIFF、GIF 或 JPEG 等格式。CCD 每次只能读取一个激发波长下的图像。对于多色荧光染料标记的芯片，自动完成更换激发和发射干涉滤光片，再次读取。由于 CCD 芯片扫描仪是同时读取整个芯片，因此不需要 X－Y 二维移动平台，扫描速度较激光共聚焦芯片扫描仪快，一般仅需 0.5～2min。

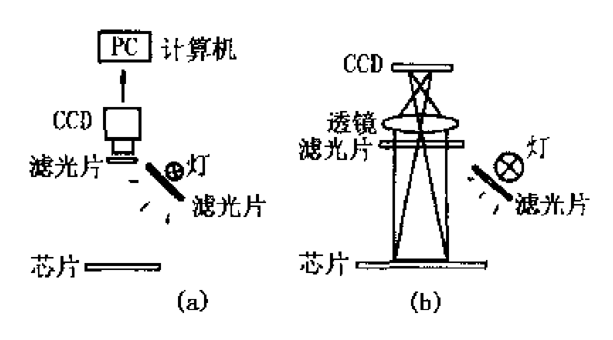


图 9－7　CCD 芯片扫描仪示意图
(a) 结构图 (b) 光路图

CCD 芯片扫描仪结构简单，价格也较激光共聚焦芯片扫描仪便宜。

9.4 DNA 芯片举例

DNA 芯片制作技术较多，玻璃为载体的芯片应用最广泛，微电子芯片技术应用电杂交，可以提高杂交特性和缩短杂交时间，很有发展前景。因此，本节主要介绍这两种芯片制作技术。

9.4.1 玻璃为载体的芯片技术

1. 多赖氨酸活化的芯片制作

(1) 芯片载体载玻片的活化

载玻璃片先用由 10%（w/v）NaOH 和 57%（v/v）乙醇组成的碱液浸泡 2h 以上，然后用双蒸馏水冲洗 5 次，载玻片浸泡在包埋液（35ml poly－L－赖氨酸（Sigma Aldrich，0.1%w/v 的水溶液）、35ml PBS 和 280ml 的双蒸馏水）中，温和振荡 1h，用大量水冲洗包埋的载玻片，45℃真空干燥 10min，室温密封保存，至少放存 2 周，使之疏水化，得到适合制备芯片的表面。

(2) DNA 芯片点样

目的 DNA 用 3×SSC 稀释，点样在处理好的载玻片上，每个点为 10nl，在 1.5×1.5cm^2 上排列 20×20 或 30×30 个点。点样后，DNA 阵列在带盖的盒子里室温干燥过夜，点样的载玻片重新水化，直至点发亮，然后 100℃快速干燥，紫外线（65Mj）照射使 DNA 与多赖氨酸化的表面交联，载玻片浸泡在 1.5% 琥珀酸酐，90% 1－甲基－2－pyrolidinone 和 43mmol/L 硼酸钠溶液中封闭未反应的表面。然后在无菌水中煮 2min，再浸入到 95% 乙醇

中，空气中晾干，室温密封保存。

(3) 探针制备

在 PCR 反应中，采用 Cy3 标记的 dNTP（200μmol/L dCTP，dGTP，dATP，40μmol/L dTTP 和 40μmol/L dUTP - Cy3），使 Cy3 掺入到 DNA 片段中，经 Sephacryl 200 旋转柱纯化，乙醇沉淀，可用于杂交。

(4) 杂交检测

带荧光素探针溶于 15μl 杂交液中（6×SSC，2×Dehardts，50%甲酰胺，0.4%SDS），90℃变性 2min，与芯片杂交，DNA 芯片用盖玻片盖住，42℃杂交过夜。杂交后，用 1×SSC/0.03%SDS 溶液室温洗载玻片，然后用 0.2×SSC 溶液洗，最后用 0.05×SSC 溶液洗载玻片，旋转干燥载片，用荧光扫描仪扫描，激发波长为 543nm。

2. PDC 活化的芯片制作

载玻片浸泡在 1%的 3-氨丙基三（甲氧基）硅烷的 95%丙酮中 2min，然后用丙酮洗 10 次，每次 5min，再在 110℃干燥 45min 后，用 0.2%1，4-亚苯基二异硫氰酸（PDC）溶液（10%吡啶/二甲基甲酰胺）浸泡 2h，然后用甲醇和丙酮洗，活化好的载玻片保存于有无水 $CaCl_2$ 的容器中。

带氨基的寡核苷酸经 HPLC 纯化后干燥，重新溶解于 100mmol/L Na_2CO_3/$NaHCO_3$（pH9.0），稀释成 2mmol/L。取 2μl 溶液点样在 PDC 活化好的载玻片上，载玻片放在 37℃含有少量水的带盖的盘子里 1～2h，然后用 1% NH_4OH 洗 1 次，无菌超纯水洗 3 次，室温下晾干。载玻片可以用于靶 DNA 的杂交检测。

在载玻片上加 10μl 杂交液（5×SSPE，0.5%SDS，20～50nmol/L 带标记的 PCR 产物），在 30℃下杂交 3h，然后用 20ml（2×SSPE，0.1%SDS）洗 15min，荧光扫描杂交结果，根据信号出现图与寡核苷酸探针布置情况，决定样品分型。

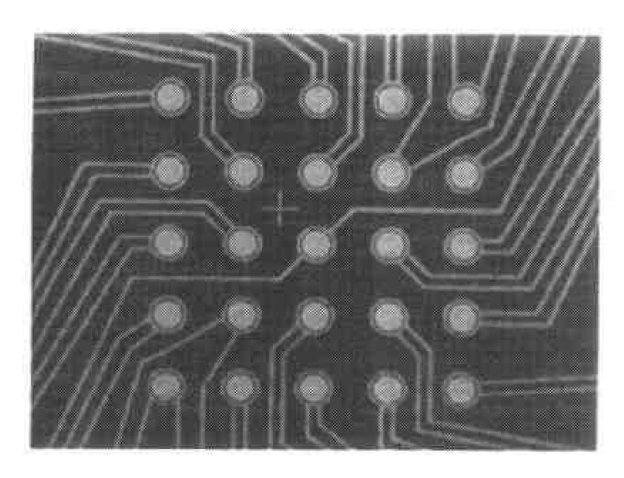
9.4.2 微电子芯片技术

1. 微电子芯片及其杂交技术的原理

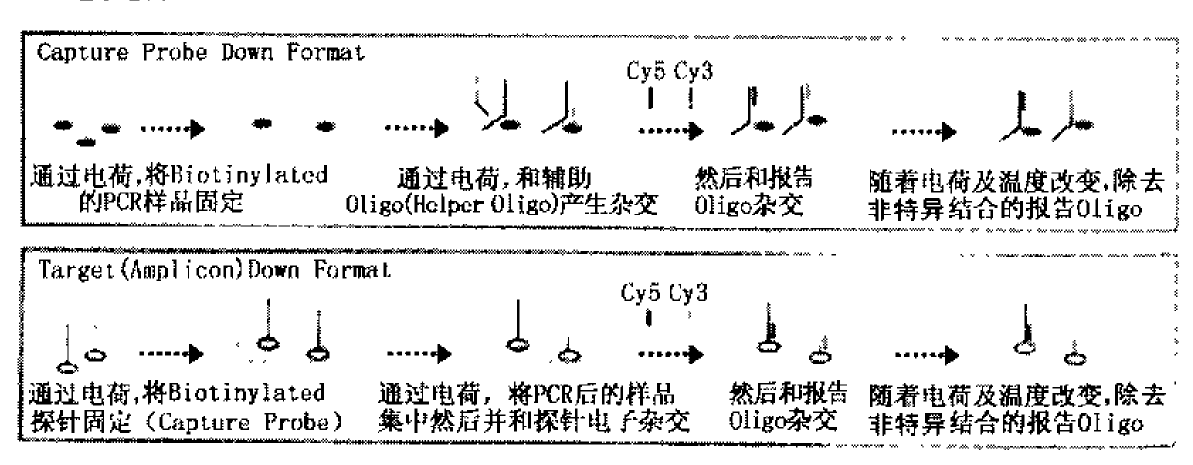
当位于待检位点下方的微电极导入一个 DC 正偏压后，DNA 捕获探针的电定位及电杂交反应就可进行。它将引起微电子阵列上待检微位点处负电荷的 DNA 的快速转运和浓缩。当核酸探针（如寡核苷酸，DNA，PCR 扩增产物，多聚核苷酸）直接与覆盖在微电极上的渗透层相连时，它就可立即被固定（共价或非共价）。接着靶 DNA 也以相似的方式被转运，浓缩，并杂交到已定位和连接的核酸探针上（图 9-9）。与被动杂交反应需要几小时相比，在微检测位点处靶 DNA 的快速浓缩，大大减少了杂交反应时间，它仅需几秒钟。

然后再逆转电场电势（负偏压），可快速除去未杂交的 DNA 分子。一旦电场调整准确，DNA 序列就从杂交的互补探针上选择性地除去。

此外，其它几个重要装置的性能和技术对促进微电子芯片的主动杂交反应也起决定性的作用。首先，起关键作用的是微电极表面的渗透层。在微芯片类仪器上，渗透层为 1～2μm 厚的琼脂糖层，该层旋转涂敷于芯片的表面。该渗透层用来保护灵敏的 DNA 杂交反应免受杂交过程中电极（铂）表面产生的电化学效应的负面影响，而且还被用作 DNA 的快速转运和随后的杂交反应。已发现组氨酸对点杂交反应特别有效。两性组氨酸溶液（接近中性 pH）电导率低，能稳定 DNA 杂交反应，并能保持较好的缓冲能力。处于典型两性



a:电子芯片



b:杂交原理

图 9-8 Nanogen 的电子芯片

状态的组氨酸溶液电导率低于 100μs/cm，而分子生物学中常用的缓冲液的传导率通常比该值大 1 000 倍左右，如 6×氯化钠/柠檬酸钠。

2. 微电子芯片制作

在 4 英寸硅晶片上涂上一层 500nm 厚的热氧化物，电极金属化的起离遮盖物（lift-off mash）由 1μm 多厚的铝光阻物（photo resist，PR）组成。再用 e 电子衍射（e-beam）工艺，分别涂上 20nm 厚的粘附层和 500nm 厚的铂层。铂由起离过程定型，即将 PR 溶于溶剂中，在铝蚀刻后用氧等离子体冲洗去除铝，lift-off 后，用等离子体增强的化学蒸镀（PECVD）法再镀上一层 2μm 厚低应力的氮化硅，PR 就被用作等离子蚀刻模具以使微电极与垫片区域相连接。接着，PR 用溶剂除去，再用氧等离子体洗净。然后芯片镀上第三层 PR，并切成小方块型。最后 PR 镀层用溶剂除去，芯片用丙酮、甲醇、异丙醇洗净，再用氮气干燥。

3. 渗透层的涂敷

（1）按照说明书，制备浓度为 25%的 Nufix 乙二醛琼脂糖水溶液；

(2) 将链亲和素溶于 250mmol/L NaCl，50mmol/L Na_3PO_4 中，使其浓度为 5mg/ml，pH 为 7.2，室温下保存；

(3) 煮沸琼脂糖溶液并冷却到 65℃，然后将琼脂糖和链亲和素混合使其终浓度分别为琼脂糖 2%、链亲和素 1mg/ml；

(4) 用旋转涂敷仪将 50μl 65℃琼脂糖涂于芯片表面（180r/min，20s），然后再在室温下将 50μl 65℃链亲合素/琼脂糖混合液涂于其上（5 000r/min，20s），接着将芯片置于 37℃烘 30min；

(5) 琼脂糖乙二醛基与链亲和素伯氨基之间的 Schiff 碱连接：用 0.2mol/L 氰硼化钠和 0.3mol/L 硼酸钠在 pH9.0，室温条件下作用 60min；

(6) 过量醛基用 0.1mol/L (pH 9) 的甘氨酸缓冲液封闭，室温作用 30min 后，将芯片用水漂洗 3 次，每次 5min，于 40℃下保存。

4. 生物素标记和荧光标记寡核苷酸的制备

合成 3′末端生物素寡核苷酸和荧光标记的 5′末端带氨基（C6 氨基接头，Glen Research）的寡核苷酸衍生物。

5. 捕捉探针的点样

(1) 将 25 个检测位点的微电子芯片置于电控工作站或 STAR 仪器系统；

(2) 芯片用 20 μl 的中性组氨酸缓冲液（50mmol/L，pH7.4）洗 3 次；

(3) 在中性组氨酸缓冲液（50mmol/L，pH7.4）中，将 20μl 50nmol/L 生物素标记的荧光对照探针定位于芯片上；

(4) 两个检测位点导入 200na/pad 正偏压 1min，四角的控制电极为负偏压；

(5) 阵列用 20μl 中性组氨酸缓冲液（50mmol/L，pH7.4）洗 5 次；

(6) 荧光下观察（594nm 激发波长，630nm 发射波长）1，1 和 5，5 位点，将发现强荧光信号，而未杂交位点仍为暗色（没有荧光信号），接着生物素标记探针以连续（每个位点定位不同的探针序列）或平行方式（多个位点定位同一探针序列）进行定位。在中性组氨酸缓冲液中（50mmol/L，pH7.4），将 20μl 50nmol/L 第一种生物素标记的寡核苷酸溶液加入到芯片上，然后在第一个待检位点处导入 200nA/pad 正偏压 1min，四角的控制电极为负偏压。将第一个捕获探针定位到该处。若要定位到多个检测位点上，电流强度（nA）也相应增大。此后，芯片再用中性组氨酸（50mmol/L，pH7.4）洗 5 次。

(7) 重复上述过程，直至所有要求的捕获探针序列都被定位到它们相应的待测位点处。最后阵列用组氨酸缓冲液（50mmol/L，pH7.4）洗 5 次，准备杂交。

6. 电子杂交

(1) 在 100μl 中性组氨酸缓冲液中（50mmol/L，pH7.4）加入 1～2μl PCR 扩增产物；

(2) 芯片用 20μl 中性组氨酸缓冲液（50mmol/L，pH7.4）预洗 3 次；

(3) 将稀释后的 PCR 扩增产物（20μl）加入到芯片上；

(4) 在待测杂交位点导入 600nA/site 正偏压 2min，四角调控电极为负偏压；

(5) 阵列再用 20μl 中性组氨酸缓冲液（50mmol/L，pH7.4）洗 5 次；

(6) 重复上述过程将其它需定位的 PCR 产物定位到其相应的待检位点处；

(7) 阵列再用 20μl 中性组氨酸缓冲液（50mmol/L，pH7.4）洗 3 次，查看芯片的荧光

图像；

（8）阵列用 20μl 20mmol/L 磷酸钠、20mmol/L Tris 缓冲液（pH9.5 特异性缓冲液）洗 10 次，将最后 20μl 留在阵列上；

（9）电子严谨度通过在匹配/错配检测位点处导入负偏压，四周调控电极导入正偏压来进行，适宜的电子严谨度电流为 400nA/site 至 1.2μA/site，脉冲频率为 0.1s off/0.1s on，循环 150 次；

（10）阵列再用 20μl 120mmol/L 磷酸钠，20nmol/L Tris 缓冲液（pH9.5）洗 3 次；

（11）然后在芯片上进行荧光检测（594nm 激发波长，630nm 发射波长）和图象分析。

微电子芯片用于 DNA 杂交分析与其它芯片相比，具有以下显著优点：

（1）电子定位过程能选择性地转运带电荷 DNA 分子；

（2）具浓缩效应的电子杂交大大增加了 DNA 的杂交速率；

（3）电子特异性控制提高了选择性并可快速鉴定出错配的碱基对。

9.5 DNA 芯片的优缺点

在 DNA 芯片技术出现以前分析研究单核苷酸的多态性和线粒体序列多态性分析多采用 PCR－SSCP，手工或自动测序方法，而所有这些又都需要经过电泳等环节，费时，成本高，不利于大规模、低消耗和自动化的要求。

应用 DNA 芯片技术进行检测可以克服上述不足。由于成千上万个寡核苷酸探针有规律地排在芯片上，能检测出无数个序列核苷酸标记物来构建每一条染色体。分辨率是第二代微卫星遗传标记物的 4～5 倍，并且对所有这些遗传标记物的分析进行一次杂交检测就可以完成，这样可以上百倍地提高工作效率，此外这种方法不会产生假阳性。可以说 DNA 芯片技术是 DNA 多态性分析中的最新和最有发展前途的方法。

DNA 芯片技术方法结果准确，特异性强。单纯的 PCR 检测可能会出现假阳性结果、或无法判断结果，通过基因芯片可以排除假阳性或非特异性 PCR 结果，从而能更准确地判断检测结果。

DNA 芯片技术方法快速，一个芯片的 DNA 不论多少个序列，只需一次杂交，在 30min 或更短的时间即可完成。DNA 芯片技术可以加速 DNA 分析的自动化。未来的法医学 DNA 检验必将采用全自动化，在 1～2h 内即可完成分析，检验过程中不涉及人工操作，大大降低了错误的发生。

但同时也应注意到现阶段 DNA 芯片技术还存在一些亟待解决的问题。如光化学合成将降低偶联率，光衍射、反射和散射等可引起寡核苷引物合成的错误等，尚需继续进行研究。

9.6 DNA 芯片的应用

DNA 芯片应用于特定基因的检测、突变检测、多态性分析、基因表达谱、基因诊断、后基因组研究、药物设计与开发、个体化医疗、身份识别和法医学、农业与环境保护。在法医 DNA 分析主要用其进行 SNP 和 STR 的分析。Lareu 报道了用微阵列技术分析了 38 个 Y 染色体 SNPs，他以载玻片为芯片载体，用 Affymetrix417 microarray 点样，在 PCR 反应过程

中掺入 Cy3 和 Cy5 荧光标记，用 microarray Affymetrix 418 扫描检测的芯片杂交结果。Reynolds 以尼龙膜为芯片载体，分析检测了 55 个 SNPs。Nanogen 设计的微电子芯片利用主动杂交进行 STR 基因座的多态性分析，不但能提高分析速度，而且能提高识别率及对降解 DNA 的分析能力。图 9－9 是 HLA DRB 芯片。

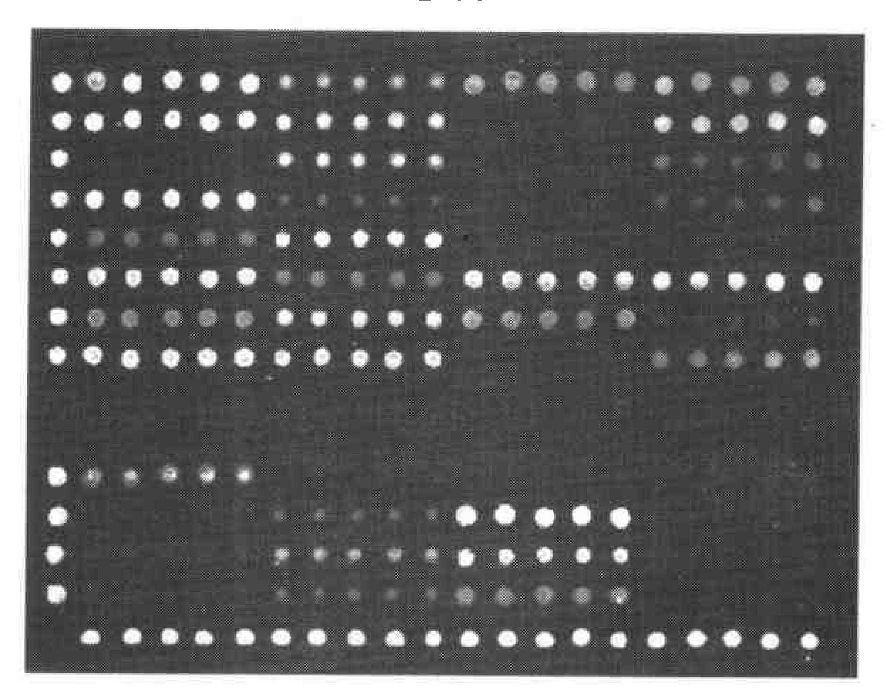


图 9－9 HLA DRB 芯片

第十章 DNA 指纹和 DNA 纹印

10.1 概述

DNA 指纹和 DNA 纹印属于限制性片段长度多态性标记，利用限制性内切酶将人类基因组 DNA 切割成不同长度的限制性片段，经电泳分离后，用特异探针杂交，检测出不同片段长度的等位基因。图谱的个体特异性取决于所用的限制性内切酶和探针选择。技术的核心是分子杂交。不同探针的 DNA 图谱原理相同，检测方法相近，只是因选用的探针和标记方法不同而略有差异。本章在操作方法上仅以辣根过氧化物酶标记 3′HVR 探针为例进行具体介绍。

10.2 DNA 指纹图—多基因座 VNTR

10.2.1 DNA 指纹—多基因座探针

人基因组 DNA 中一系列小卫星重复单位的核心序列有高度的同源性，因此，小卫星序列可以交互杂交，如采用核心序列多聚体作为探针，在低强度杂交条件下，只要与探针序列同源性达到能稳定杂交的所有小卫星片段均可同时被检出，一次可检出多个基因座，这类小卫星核心序列多聚体探针叫多基因座探针（multi - locus probe），如 33.6、33.15、Myo、MZ13 等属于此类探针（表 10 - 1）。用多基因座探针检测得到的 DNA 图谱称为 DNA 指纹图。

表 10 - 1 部分多基因座探针

探针名称	限制酶	平均片段数	偶合率
33.15	HinfⅠ或 HaeⅢ	15	3.0×10^{-11}
33.6	HinfⅠ或 HaeⅢ	11	3.0×10^{-9}
α - 珠蛋白 - 3' - HVR	HinfⅠ或 HaeⅢ	20	3.7×10^{-14}
Myo	HinfⅠ	14	4.0×10^{-9}
JL02	HinfⅠ	16	6.60×10^{-15}

10.2.2 DNA 指纹图谱特征

多位点 DNA 指纹图利用小卫星序列具有同源性的特征，在低强度杂交条件下，使 DNA 探针与不同染色体上 VNTR 基因座的靶片段杂交。因此 DNA 指纹图含有许多片段，在 3～23kb 范围内，有 10～30 条片段，图谱极其复杂且携带大量的遗传信息，具有人类遗传标记的各种特征。

1. 孟德尔遗传

通过大量的家系调查，发现 DNA 指纹图中大多数谱带，尤其是 4kb 以上的大片段以均等机会由亲代传递给子代，子代指纹图谱的所有谱带都可在双亲的指纹图谱中找到，片

段的传递符合按孟德尔遗传规律，但子代的谱带不是简单的各一半来自父亲、母亲。来自母亲的谱带数可能与来自父亲的谱带数不同。未发现任何一条谱带是特异地从父亲或母亲传递给子女的，排除性连锁遗传的可能性。此外，DNA 指纹图中各谱带传递是独立的，谱带间没有连锁关系。只有当生殖细胞发生突变，才有可能导致子代个体 DNA 指纹图的改变，出现双亲不存在的陌生谱带。

2. 组织同一性

DNA 指纹图无论在体细胞还是在生殖细胞中都是稳定的，同一个体的不同组织的 DNA 指纹图都是相同的，无组织间差异。同一个体的血液、精斑、毛发、肌肉、脑组织、骨等的 DNA 指纹图谱完全一致。离体细胞培养的细胞株仍保持与供体一致的 DNA 指纹图。

但有两种情况例外：一种是不同组织之间甲基化程度不一样，当使用对甲基化敏感的限制酶如 HinfⅠ时，不能识别和切割甲基化的酶识别点，导致同一个体不同组织 DNA 指纹图不一致。但改用耐甲基化限制酶如 HaeⅢ时，这种差异就会消失；再一种就是组织细胞的癌变。由于癌变细胞染色体丢失 DNA 或异常扩增，导致某一 DNA 谱带丢失或密度增强，或出现额外带，使癌变组织与正常组织 DNA 指纹图不同。

3. 高度的个体特异性

通过对 DNA 指纹图的统计计算表明，平均匹配概率在 10^{-12}以下，也就是说在 100 亿个人中，方可找到两个 DNA 指纹图谱完全相同的个体，足以说明 DNA 指纹图的高度个体特异性。惟一例外的是同卵双生子，他们的 DNA 指纹图完全一致。

4. 突变率

分析父、子、母的真三联 DNA 指纹图，发现子代 DNA 指纹图中出现父、母均没有的“陌生带”，这是突变形成的新片段。早期 Jeffreys 等用探针 33.6 和 33.15 分析一个亚洲印第安人真实家系 4 代 54 个个体时，发现平均每 300 条子代片段中出现 1 条陌生带。对 1419 例真实家系调查结果分析，33.6 探针的突变率为 0.52%；33.15 探针的突变率为 1.1%（表 10－2）。在 1419 个真实家系子代中，出现 1 条陌生带的有 82 例（26.9%），出现 2 条陌生带的有 7 例（1.2%）。Jeffreys 以及 Smith 等人对真实家系的调查发现，子代 DNA 指纹图中含有 1 条或 2 条陌生带的情况较常见。因此，在应用 DNA 指纹图进行亲子鉴定时，遇到存在 1 条或 2 条陌生带时仍不能否定亲子关系，应参考其他方法如 PCR、单基因座 DNA 纹印图的检验结果综合分析。在子代 DNA 指纹图出现 3 条或 3 条以上陌生带的案子中，可以排除亲子关系。

突变片段出现的主要原因是细胞在减数分裂过程中出现了基因交换与重组。VNTR 基因座为高变区，多数小卫星序列本身就是基因组 DNA 内的重组热点。VNTR 基因座的重组与交换同样也是等位基因片段的主要原因，突变率可高达 28.1%，突变率过高对个体识别影响较小，但对亲子鉴定不利。

表 10－2 家系分析中子代DNA指纹图与父母DNA指纹图关系的调查

	33.6	33.15
父方谱带平均数（±SD）	8.08±2.47	7.58±2.22
母方谱带平均数（±SD）	8.11±2.47	7.67±2.29
公共谱带平均数（±SD）	2.48±1.58	2.22±1.57
陌生带数目	0.098	0.195
突变率/子代谱带	0.0052	0.0113
(95%可信限范围)	(0.0044～0.0062)	(0.0098～0.0124)

10.2.3 DNA指纹图匹配率的计算

多基因座DNA指纹图是多基因座探针在低强度杂交条件下与多个VNTR基因座的多态片段杂交形成的。在人为地控制杂交条件下同时检测多个VNTR基因座，VNTR多态性片段与探针序列会有一定的错配率。1条片段，可能来自某纯合子基因座上相同长度的基因，也可能是来自2个或2个以上基因座的相同长度基因。在DNA指纹图上无法判定某片段来自哪一个多态基因座；无法确定各基因座的基因型，也不能通过群体调查确定等位基因及基因型频率。多基因座DNA指纹图上数10条谱带只是小卫星探针应检出等位基因的一小部分，目前仍不能准确估计有多少基因座的多态性片段被漏检。多基因座DNA指纹图的分析方法、统计学计算原则与一般遗传标记系统明显不同。

虽然DNA指纹图十分复杂，但所有片段均来自VNTR基因座，每条等位基因片段的遗传应符合孟德尔遗传法则，这是多基因座DNA指纹图能成为人类遗传标记系统的先决条件。假设指纹图中的片段不论大小，每一片段在群体中平均频率一致，无关个体间共有一个片段的概率一致；假设同一片段长度区内，不同个体间检出的片段数相同，没有差异，并且不出现因突变形成的新片段；假设由同一探针检出的多个VNTR基因座间无基因连锁关系，统计学上具有独立性。当上述三个假设成立时，则可以通过群体调查研究人群中每个片段的平均频率（q）和无关个体共有一个片段的平均概率（x）。x值是目前计算指纹图偶合概率的重要基础数据。由平均概率计算出两个比对样品之间的相关指数或偶合率确定他们的关系。

在多基因座DNA指纹图中，每一条带可以认为是来自某个具有一带A和a的基因座。设A的频率为q，a的频率则为1－q。在处于遗传平衡的群体，A谱带出现的频率x_i应是AA和Aa基因型频率之和，即$x_i = q^2 + 2q(1-q) = 2q - q^2$。因为小卫星DNA高度变异，每一条带（等位基因）频率很低，所以q^2带可忽略不计，则$x_i \approx 2q$。该式表明了群体中共有一条带的概率X和片段基因频率之间的数量关系。x_i即一个体的A谱带在另一个体中出现的概率。

为了计算方便，一般按分子量标准如λDNA/HindⅢ＋EcoRⅠ划分成几个区，先统计比较每一区域内的总谱带数N_i和既在A个体又在B个体出现的谱带数目以计算平均概率与偶合率。例如，调查6个个体，每个个体最多能检出5个片段，各个体情况如表10－3所示。在6个人中，具有2.3kb片段的有3人（片段共有概率为3/6＝0.50），具有2.9kb片段的有4人（片段共有概率为4/6＝0.67），4.1kb片段的有2人（片段共有概率为2/6＝0.33），8kb片段的为5人（片段共有概率为5/6＝0.83），11.7kb片段有3人（片段共有概

率为 3/6 = 0.50)，二个人共有一个片段的平均概率（X）为上述 5 个片段共有概率的平均值，即 X = （0.50 + 0.67 + 0.33 + 0.83 + 0.50）÷ 5 = 0.57。将此值代入 X = 2q = 0.57，则平均基因频率 q = 0.285。6 个个体总片段数为 17，每个个体平均片段数 n = 17 ÷ 6 = 2.83，

个体平均片段数为 n = 2.83 ± 0.98。根据上述调查结果，该群体中两无关个体的 DNA 指纹图完全一致（所有片段均共有）的偶合率为 $X^n = 0.57^{2.83} = 0.20$。

表 10-3　多基因座 DNA 指纹图的群体片段调查实例

个体编号 片段（kb）	1	2	3	4	5	6	片段共有概率
2.3	+		+	+			0.50
2.9		+		+	+	+	0.67
4.1	+		+				0.33
8.2	+	+	+		+	+	0.83
11.7	+	+	+	+			0.50
总计	4	2	4	3	2	2	0.57

10.2.4 几种 DNA 指纹图介绍

国内从 1989 年开始有多基因座探针 DNA 指纹图在法医学上的应用研究报道，如“α-珠蛋白 3′HVR”（李伯龄等 1989）、“MYO”（姜先华等 1990）、“$(CAC)_5$/ $(GTG)_5$”（蓝翎等 1993）、人工合成的寡核苷酸探针“JL-02”（季安全等 1997）、以及“33.15”和“33.6”（唐晖等 1998）等等。所建立的多基因座 DNA 探针检测 DNA 指纹图方法，其两个无关个体间出现同一条带的平均相关机率为 0.2～0.25，两随机个体谱带完全相同的机率为 $2.0 \times 10^{-9} \sim 2.0 \times 10^{-15}$。

1.α-珠蛋白 3′HVR 探针指纹图

α-珠蛋白 3′HVR 探针最初是由 Higgs 等人分离得到的，它由 1 个 17 bp 的随机重复单位组成，寡核苷酸序列为 5′-GNGGGG [N] ACAG-3′。α-珠蛋白 3′HVR 探针在高强度杂交条件下可作为单基因座探针使用，在低强度杂交条件下作为多基因座探针使用。

用 α-珠蛋白-3′HVR 探针/限制酶 HinfⅠ对 200 名无关汉族个体进行 DNA 多基因座指纹图分析，平均每个体片段数 16.8 条，同位素 ^{32}P 标记的探针杂交所得图谱的片段平均概率 0.198（表 10-4），HRP 标记的探针检测片段共有概率为 0.21。由此计算两汉族无关个体 DNA 指纹图相同的偶合概率 $P = 0.198^{16.8} = 1.5 \times 10^{-12}$（图 10-1）。

表 10-4　3′HVR/HinfⅠDNA 指纹群体调查参数(李伯龄等)

探针	片段长度(kb)	个体片段均数(± s)	平均片段概率(X)	平均基因频率(q)
α-珠蛋白 3′HVR	9.6～23	1.3 ± 1.0	0.135	0.068
	6.4～9.6	3.1 ± 1.3	0.173	0.086
	4.4～6.4	6.1 ± 1.5	0.231	0.116
	3.0～4.4	6.3 ± 1.5	0.252	0.128
	3.0～23	16.8 ± 1.3	0.198	0.099

2. 33.6 探针和 33.15 探针指纹图

33.15 和 33.6 探针是揭示 DNA 指纹图的第一批探针，1985 年 Jeffreys 在研究人肌红蛋

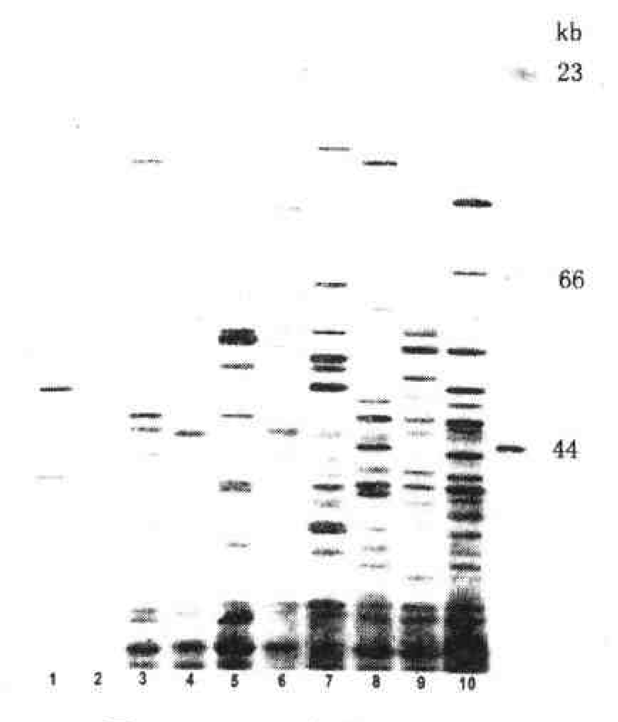


图 10-1　3'HVR 探针 DNA 指纹图

白基因时发现的。当时他们发现在第一个内含子中小卫星重复序列与其他 3 个已知小卫星（胰岛素基因、α-珠蛋白基因和 C-Ha-ras 基因）在序列上有一定的同源性。他们便将肌红蛋白基因内含子小卫星 DNA 重复单位头尾相连克隆至 pUC13 载体上，再从阳性重组子中分离纯化由 23 个重复单位构成的插入片段，以该片段（pAV33.7）为探针在低强度杂交条件下检测人类基因组 DNA 时，不同小卫星 DNA 便可同时被检出，他们用该探针筛选人类基因组 DNA 文库 λL47.1，从 3×10^5 个重组子中筛选出 40 个阳性克隆，随机取 8 个克隆进行 Southern 印迹杂交和序列分析，结果表明阳性克隆的重复单位间有一段序列是基本相同的，Jeffreys 称之为核心序列。阳性克隆间重复单位长度呈倍数关系，如：λ33.5，33.15 为 16 bp，λ33.3，33.6，33.10，33.11 为 33bp，λ33.1，33.4 为 64 bp，这表明 pAV33.7 探针所检测的小卫星 DNA 不仅依赖于核心序列的存在，同时还依赖于重复单位核心序列与小卫星间交互杂交形成杂合分子时的相对关系。他们所筛选出的 λ33.15 克隆的插入片段（探针 33.15）是有由 16 bp 的核心序列前后重复 29 次所构成，以此为探针在低强度杂交条件下进行限制性片段长度多态性分析，所得图谱比 pAV33.7 探针的图谱更为复杂，谱带数更多，具有高度的个体特异性。另外，还有 33.6 探针符合多基因座探针。

33.6 探针核心序列为：$(AGGGCTGGAGG)_{18}$

3.15 探针核心序列为：$(ACAGGTGGGCAGGTGG)_{29}$

33.15 探针检测 Hinf Ⅰ 或 Hae Ⅲ 消化基因组 DNA 的产物。在 4～20 kb 范围内平均能检测到 15 条带，33.6 探针平均能检测到 11 条带。33.15 探针检测无关个体 DNA 指纹图完全相同的相关机率达 3.0×10^{-11}，33.6 探针检测无关个体 DNA 指纹图完全相同的相关机率达 3.0×10^{-9}，两个探针加在一起的相关机率为 5.0×10^{-19}（表 10-5）。

表 10-5 33.6 和 33.15 探针的群体调查参数（Jeffreys）

探针	片段长度（kb）	个体平均片段数目（±SD）	平均片段频率（X）	平均等位基因频率（q）
33.6	10~20	2.8±1.0	0.11	0.06
	6~10	5.1±1.3	0.18	0.09
	4~6	5.9±1.6	0.28	0.14
33.15	10~20	2.9±1.0	0.08	0.04
	6~10	5.1±1.1	0.27	0.14
	4~6	6.7±1.2	0.20	0.10

3.Myo 探针指纹图

Myo 探针是参照人肌红蛋白基因小卫星 33bp 重复序列合成。以 33 个碱基为一个重复单位连接使之重复 15 次，长度大约为 0.5kb，然后将此 DNA 片段克隆至 pUC19 中，其结构为：

GAC CGA GGT CTA AAG CTG GAG GTG GGC AGG AAG

CTG GCT CCA GAT TTC GAC CTC CAC CCG TCC TTC

将 Myo/pUC19 质粒转化至大肠杆菌 HB101 中，然后进行质粒的扩增和提取。用 Myo 探针检测 Hinf Ⅰ或 Hae Ⅲ消化的基因组 DNA 产物，可获得与 33.6 和 33.15 探针类似的多基因座 DNA 指纹图。在 2.3~23 kb 范围个体片段数平均为 14 条，两个无关个体共有一条带的平均概率为 0.215，无关个体 DNA 指纹图完全相同的相关机率为 4.0×10^{-9}（表 10-6）。

表 10-6 Myo 探针群体调查参数（姜先华等）

探针	片段长度（kb）	个体平均片段数目（±SD）	平均片段频率（X）	平均等位基因频率（q）
Myo	23~9.6	1.5±1.2	0.184	0.074
	9.6~6.6	2.28±0.83	0.195	0.0975
	6.6~4.4	2.94±1.21	0.226	0.113
	4.4~2.3	6.83±2.94	0.292	0.146

4. 寡核苷酸多基因座探针 $(CAC)_5$/ $(GTG)_5$ 指纹图

人类基因组 DNA 中散在分布着一些类似于小卫星的简单重复 DNA，如（GACA）n、（AGAT）n、（AATC）n、（CA）n 等，在 DNA 多态基因座分类之中被称为微卫星（microsatellite），它与小卫星 DNA 形成原因、分布及多态性特征上极为相似，只是重复单位简单，由少数碱基对组成。相应检测这类 VNTR 基因座的探针组成也十分简单，如 $(CAC)_5$/ $(GTG)_5$ 探针，比较容易由人工合成获得，并称之为寡核苷酸多基因座探针。

$(CAC)_5$/ $(GTG)_5$ 探针检测 Hinf Ⅰ 酶切基因组 DNA 图谱，在 4kb 以上平均有 13 条谱带，无关个体之间谱带差异显著（图 10-2）。任意两无关个体的偶合概率为 3.8×10^{-10}（表 10-7）。

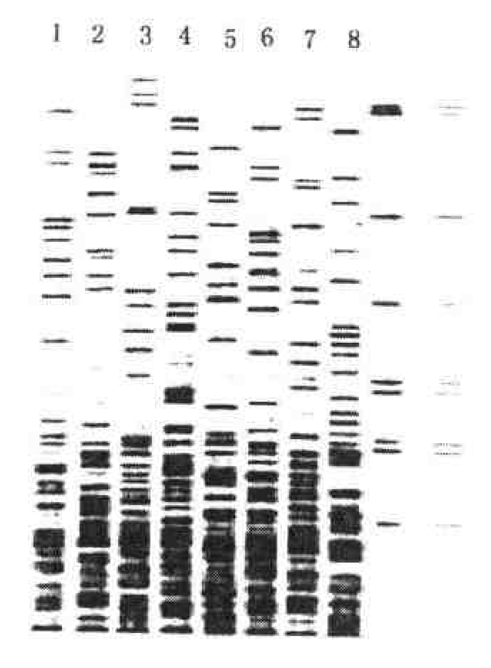


图 10－2 $(CAC)_5$/ $(GTG)_5$ 探针 DNA 指纹图

表 10－7 (CAC) 5/ (GTG) 5 探针群体调查参数（兰钢等）

探针	片段长度（kb）	个体平均片段数目（± SD）	平均片段频率（X）	平均等位基因频率（q）
$(CAC)_5$/$(GTG)_5$	24～9.4	4.4 ± 1.0	0.10	0.05
	9.4～6.5	3.7 ± 1.4	0.21	0.11
	6.5～4.0	4.8 ± 1.3	0.30	0.16
	4.0～24.0	12.9 ± 1.1	0.20	0.11

5. MZ1.3 探针指纹图

MZ1.3 多基因座 VNTR 探针也是从人基因组 DNA 中筛选的小卫星探针。用限制性内切酶 Sau3A 部分消化人类基因组 DNA，将酶切片段插入到 λ－EMBL 噬菌体 DNA 重组构建人基因文库。然后用完整的 M13mp18 噬菌体 DNA 为探针筛选阳性克隆，一共筛选出 6 个阳性克隆。将其中一个阳性克隆片段连接到质粒 pUC18 中，组成一个重组子。将 MZ1.3/pUC18 质粒转化至大肠杆菌 HB101 中，然后进行质粒的扩增和提取，最终得到 MZ1.3 探针。MZ1.3 探针是一个来源于 λ 基因克隆的用 Pst Ⅰ/ Bam H Ⅰ 水酶解的 1.85 kb 大小的片段，含有一个小卫星序列，核心序列为 27bp：GC$^{T}_{A}$GGGGTGGTGGTCGGTCTGGCCGGC 共 47 个拷贝。MZ1.3 与噬菌体 M13 的 15bp 重复单位有 53%～73% 的同源性。

在 4.3～23 kb 之间平均有 18 条带，两个无关个体间同时出现同一条带的平均概率为 0.205，无关个体 DNA 指纹图完全相同的相关机率为 6.0×10^{-12}。

6. JL－02 探针

寡核苷酸探针核心序列为 $(NGG)_5$ $(GAGGTG)_3$；用外展型扩增反应制成长度更长的

探针。

JL－02 探针检测的 HinfⅠ酶切基因组 DNA 的指纹图的大片段谱带较多。在 4.3 kb 以上的谱带中，每个个体平均检出 16 条谱带，任意两个无关个体的偶合概率为 6.6×10^{-15}（表 10－8）。

表 10－8 JL－02 探针群体调查参数（季安全等）

探针	片段长度（kb）	个体平均片段数目（±SD）	平均片段频率（X）	平均等位基因频率（q）
Jl－02	23～9.6	4.12±1.71	0.044	0.022
	9.6～6.4	5.28±1.36	0.166	0.083
	6.4～4.4	6.84±1.34	0.222	0.111

10.2.5 DNA 指纹的应用

DNA 指纹技术的建立被迅速应用到法医学的个人识别和亲子鉴定中，对生物物证的鉴定起到了巨大的推动作用。

1. 个人识别

多基因座 DNA 指纹图提供的信息量大，除同卵双生同胞外，任何两个无关个体的指纹图谱完全不同。在同一自显影照片中两份生物检材相互比对，如果两生物检材的谱带数目与位置不同，则可肯定两份检材不属于同一个体；当两份指纹图的谱带位置与数目均一致，称之为匹配（match），可肯定两份检材来自同一个体。目前应用的数个多基因座探针在 3.0～20kb 范围内均可检测出 10～30 条多态片段，群体中片段共有概率均在 0.25 以下。图 10－3 为一例强奸案的多基因座 DNA 指纹图，1 号现场精斑 DNA 与 2 号嫌疑人血液 DNA 经限制酶 HinfⅠ酶切，α－珠蛋白－3′－HVR 探针杂交，自显影图片中显示 3kb 以上的有 13 条，按共有一个片段平均概率 X＝0.21 计算，两者来源于同一个体的偶合概率为 $0.22^{13}=2.1\times10^{-12}$，偶合概率接近 0，表明某无关的两个体被误认为同一个体的可能性几乎不存在，可以认定精斑为嫌疑人所留。

在检验实际案件检材时应考虑到，现场物证检材情况往往比较复杂，可能量少或受到污染或 DNA 出现降解，指纹图中的谱带数目可能比新鲜的比对样品少，所以不能单纯凭谱带数目不同作出排除结论，关键要看两者已出现谱带数和位置是否完全对上。例如图 10－3 中 1 号精斑比 2 号嫌疑人（箭头所指）的谱带少。两者仅仅是谱带数不同，谱带少的检材中所有的带均可在谱带多的检材中找到，通过计算偶合率，偶合率很小（$<10^{-7}$），仍可以考虑作出肯定结论。或结合其他检验结果作出结论。

多基因座 DNA 指纹技术比较复杂，必须有严格的操作程序及严格的质量保证体系。生物检材的个人同一认定，必须在同一条件下进行 DNA 指纹检验，其结果才有可比性。指纹图谱带完全匹配的假阳性结果不易产生，但由于检材条件和操作中各种因素干扰，有可能出现假排除现象。假排除的原因主要是来自生物检材复杂因素的干扰，检材种类多样性，保存条件不一，增加了 DNA 指纹检验的难度，不能保证获得完整的 DNA 指纹图。

现场采集提取的检材，因含盐分较多，容易发生阳极漂移，DNA 多态片段较新鲜血液的移动快。通过检查每个样品均拥有的无多态性 DNA 片段的位置是否一致，来确定是



图10 3 DNA指纹分析强奸案件结果图

否有阳极漂移存在，如果这些无多态性DNA片段匹配不上，整体向前移动，说明有阳极漂移存在，另外，DNA量不同，也会出现漂移现象。结果分析时，必须十分小心。

基因突变、基因组DNA的酶切不完全、或检材受其他DNA的污染，与多基因座探针形成非特异性杂交，DNA指纹图中产生多余谱带。产生1~2条谱带并不影响个体同一认定，但DNA酶切不完全产生多余带的数目与位置则难以估计，因此在电泳前应进行对酶切程度的检测。物理、化学或生物因素引起DNA分子的损伤，导致DNA的降解。降解DNA分子不能作多基因座DNA指纹鉴定，部分降解DNA检材也会引起指纹图中额外谱带。

样品为几个个体的混合物时，DNA指纹图中情况将更为复杂，当与某一嫌疑人的DNA指纹图相比较时就会发现有几条相同的带，又有几条不同的带。在这种情况下，应尽量收集现场各部位精斑，不同部位精斑分开检验，对较大面积的斑迹按精子量的多少分取2~3个部位分开检验，这样就有可能遇到某一个人的精斑或主要为某一个人的并有少量其他人的精斑，便于DNA指纹的判定。

部分检材DNA指纹图的片段减少，有2个常见原因：一是基因组DNA部分甲基化，可换用HaeⅢ限制酶；二是DNA用量少，可增加检材DNA量加以克服。

2. 亲权鉴定

DNA指纹图的谱带遵循孟德尔遗传定律传递给子代。子代指纹图中所有谱带均应出现在父亲或母亲的指纹图中。比较同一自显影片的孩子、假定父母的指纹图谱，分析三者间的关系。如图10-4所示，先在孩子指纹图中找出与母亲共有的谱带，剩下的谱带均为非母亲带（设有n条），如果没有突变带出现，所有n条非母亲带均应在生父指纹图中找到。当孩子的n条非母亲带中至少有3条以上不能在假定父中找到，则排除假定父与孩子的亲子关系；如果孩子的谱带均能在假设父母中找到，或只有1~2条不能在假定父母指纹图中找到，根据无关个体共有1条带概率，计算他们的相关机率。例如，子的非母带有m条（$n-2 \leq m \leq n$）在假定父中找到，计算他们的相关机率X^m（X为无关个体共有1条

谱带的概率)，认定是否有父子关系。图10－4为一亲子鉴定的指纹图。

图10－4 多基因座DNA指纹图在亲子鉴定中分析模式图

A：认定亲子关系例子；B：否定亲子关系例子。1：自显影照片原图；2：在孩子图谱中标记出来自母亲方的带；3：确定剩余的带是否来自于假定的父亲？

DNA指纹图的高突变率也是一个特点，亲子鉴定时发现有1～2条陌生带出现，下结论应慎重，要结合其他探针、DNA纹印图或PCR方法等检验作出结论。

10.2.6 DNA指纹的局限性与潜在的问题

多基因座DNA指纹图一次检测获得最多遗传信息，其高度个体特异性是其他任何遗传标记系统无法比拟的，但在法医学应用中，也有其局限性，存在潜在问题。

1. 灵敏度低，检材用量大：一次检测多基因座DNA指纹图至少需要0.5～1.0μg DNA，相当于50μl的血液（痕）、5μl精液、0.5～1.0μg组织或15根毛发。一般案件的检材难以满足此条件。

2. 对基因组DNA的质量要求高：提取的检材DNA分子至少在20kb以上，腐败检材或降解DNA无法进行多基因座DNA指纹图检验；并且检材DNA中不能含有抑制限制性内切酶活性的抑制剂。如果样品DNA中含有色素等杂质，必须先进行DNA纯化，否则抑制内切酶的消化作用，得不到DNA指纹图。

3. 多基因座DNA探针没有种属特异性：实验证明，探针33.15和33.6不仅与人DNA有同源性，而且对狗、猫、鸡、鸟甚至鱼类的DNA也有同源性，可以发生杂交。同样，α

-珠蛋白-3′-HVR探针与牛、马、狗、鸡、鹿、鱼等DNA也有同源性；MZ1.3探针是用噬菌体DNA筛选出来的，与常见野生和家养动物DNA均有相当高的同源性。常用的多基因座探针均能杂交检测出动物DNA。当物证检材中污染有动物DNA时，可出现多余的额外带，严重干扰对物证检材的检验，导致错误的结论。

4. 高突变率：在真实家系的子代存在1~2条陌生带（突变带），陌生带的出现可能导致假排除。因此出现陌生带时，判断结果要慎重，必要时应同时检测2个多基因座DNA指纹图或进行其他系统的检验，避免误判。

5. 检验操作繁琐，实验环节多，实验周期较长（5~7天）。

6. 存在统计学问题，指纹图谱带数多，而且呈连续分布，测出的众多等位基因没有明确的基因定位，个体同一性概率、亲子关系概率及群体数据的统计学计算方法的理论尚待完善。

7. 电泳凝胶的变形和谱带的漂移等将不可避免带来误差，因此不是同一块凝胶的分析结果不能比对。这也就造成在以往检验过的案件中，如果嫌疑人被排除，在以后重新找到嫌疑人，一旦原有检材已用完将不能再进行比对检验。

8. 不能解决来自2个以上个体混合检材的个体识别问题。

10.3 DNA纹印图—单基因座VNTR

10.3.1 单基因座探针

1. 探针特点

单基因座探针（single locus probe）指在高强度条件下，只检测出单个基因座多态片段的探针，因此又称为基因座特异性探针（locus specific probe），习惯称单位点探针。由单基因座探针检测得到的图谱称为DNA纹印图。单基因座探针检出的多态性片段在杂合子个体中可检出2条带，即该基因座的2个片段长度等位基因，纯合子为一条带。

单基因座探针多从基因组DNA中筛选，属小卫星探针。筛选评估探针指标除高特异性要求外，还包括：探针所检测基因座的染色体位置，多态片段的长度范围，等位基因数及在群体中频率分布，基因座杂合度、突变率及探针的交叉反应等。表10-9列出了部分单基因座探针所检测的基因座，使用的限制性内切酶，杂合度和片段长度范围等供选用参考。

表10-9 部分单基因座探针

探针名称	基因座	限制酶	杂合度	片段长度（kb）
3′HVR	D16S85	Pst Ⅰ	0.90	0.5~6.0
		Hae Ⅲ	0.90	0.5~6.0
		Pvu Ⅱ	0.93	0.5~6.0
g3	D7S22	Hinf Ⅰ	0.98	1.5~22.0
L45	D5S61	Msp Ⅰ	0.76	2.8~7.4
L892	D3S17	Taq Ⅰ	0.71	2~22.0
MS1	D1S7	Hinf Ⅰ	0.99	3.5~16.0

续表

MS43	D12S11	HinfⅠ	0.97	3.5～13.0
MS31	D7S21	HinfⅠ	0.98	3.5～13.0
pAC601	D14S1	PstⅠ	0.95	3～19.0
pAC601	DXYS14	PstⅠ	0.98	1～10.0
pAC256	D17S79	PstⅠ	0.85	2.5～6.0
pAC229	D14S13	PstⅠ	0.95	2～18.0
pAC404	D18S27	PstⅠ	0.95	2～8.0
pCMM101	D14S13	HaeⅢ	0.95	1.0～3.5
pCMM86	D17S74	HaeⅢ	0.98	－
pL336	D1S47	PstⅠ	0.90	2～12.0
		PsaⅠ	0.88	4～12.0
pL427－4	D21S112	PstⅠ	0.94	1.4～ 3.4
		AluⅠ	0.94	0.5～ 2.5
pL355－8	D20S15	PstⅠ	0.66	6～ 8.0
		HaeⅢ	0.66	0.8～2.3
pL365－1	D11S129	PstⅠ	0.66	1.5～3.0
		HaeⅢ	0.66	0.8～2.6
		AluⅠ	0.66	0.5～2.5
pL159－1	D18S17	PstⅠ	0.74	5～7.0
pS194	D7S107	PstⅠ	0.85	5.21～12.42
		TaqⅠ	0.80	7.9～14.0
pYNH24	D2S44	HaeⅢ	0.97	0.6～5.0
Sli103	D2S44	PstⅠ	0.90	6.5～17.0

10.3.2 DNA 纹印图谱的基本特征

1. 高多态性与组织同一性

对于 VNTR 基因座，理论上相差 1 个重复单位就构成 1 个片段长度等位基因，因此 VNTR 基因座上的等位基因数之多、杂合度之高是任何常规血型系统无法比拟的。一个高度多态性 VNTR 基因座上的等位基因少则十几个，多达数百个，尤其是高变区内的基因座，都呈高度杂合状态，有的杂合度几乎接近 100%。例如 D16S85 基因座的重复单位为 17bp，选用探针和限制性内切酶 PstⅠ检出多态性片段范围为 0.5～6kb，该基因座理论等位基因数为（6 000÷17）≈353 个。又如 D1S7 基因座重复单位长 9bp，片段长度范围 1～20kb，按 9bp 的增量确定等位基因，D1S7 基因座的等位基因数将有 2 000 多个。假设每个等位基因频率相等（1/2 000），则该基因座杂合度等于 1－2 000/（2 000×1999）＝0.9996。理论基因型数为 C_{2000}^2＋2 000＝2 001 000，表明至少 200 万人中才有一人拥有相同的基因型。表 10－10 为根据中国北方汉族群体在 D2S44、D17S70、D14S13、D14S1、DXYS14 和 D18S27 6 个基因座的基因频率分布和计算的 6 个单基因座纹印图的个人识别机率和非父排除率。6 个基因座的累积个体识别机率和累积非父排除率几乎达到了 100%。

单基因座 DNA 纹印图是法医学个人识别、亲子鉴定的有力工具。

与多基因座 DNA 指纹一样，各种组织具有同一性，可以根据不同组织进行个体同一认定。

表 10-10　汉族群体单基因座纹印图 DP 和 P_E 值（刘明俊）

探针	基因座	H	P_E	DP
pAC255	D2S44	0.75	0.8407	0.9880
pAC256	D17S79	0.76	0.7691	0.9766
pAC299	D14S13	0.92	0.9792	0.9996
pAC061	D14S1	0.89	0.9049	0.9961
pAC225	DXYS14		0.9400	0.9973
pAC404	D18S27	0.76	0.8177	0.9838

2. 孟德尔遗传规律

家系调查证明，单基因座 DNA 纹印图中片段的传递遵循孟德尔遗传规律，子代的 2 条带 1 条来自于母亲，另 1 条来自于父亲。图 10-5 是以 PLMJ14 探针对一个家系的分析情况。



图 10-5　PLMJ14 探针对一家系 RFLPs 分析结果

3. 较高的突变率

减数分裂中出现了不等交换和重组，使等位基因的重复单位数目增加或减少，形成新的等位基因，在家系分析时孩子 DNA 图谱中出现不能在生父和生母找到的谱带，即陌生带。高突变率是 VNTR 基因座的特点之一（表 10-11）。DNA 纹印图能直观地揭示某一基因座的突变情况，新的突变对个体的同一认定不会有太大的影响，但在亲权鉴定中新出现的突变片段可能会导致排除生父的错误。要求单基因座 DNA 纹印的突变率不得大于 0.2%。对于个别突变率较高的探针，如 D1S7 突变率大于 0.0380，不适合于亲权鉴定。

表 10-11 单基因座 DNA 纹印图的突变率（白人群体）

基因座	探针	突变率
D14S1	pAC061	0~0.0004
D2S44	pAC255	0~0.0005
HRAS	Ha-ras	0~0.0003
D17S79	pAC256	0~0.0005
DXYS14	pAC225	0.0010~0.0030
D5S43	MS8	0~0.0030
D12S11	MS43	0~0.0030
D7S22	g3	0.0006~0.0090
D7S21	MS31	0.0030~0.0150
D1S7	MS1	0.0380~0.0750

4. 种属特异性

单基因座探针均经过严格的筛选，在高强度杂交条件下，它不仅具有基因座特异性，而且也具有高度的种属特异性，与其他种属来源 DNA（包括细菌 DNA）不杂交，因此即使混有其他动物 DNA 的污染检材，其 DNA 纹印图谱亦不受干扰。

10.3.3 DNA **纹印图的基因频率**

1.DNA 纹印等位基因特点

DNA 纹印的等位基因数多，一个基因座上的等位基因数可达数十个，甚至数百个，片段大小跨度大。为了顾全大小片段分离效果，常用 0.6%~1.5% 琼脂糖凝胶分离酶切片段，无法分辨所有仅差一个重复单位的两个 DNA 片段，分辨率有限。琼脂糖凝胶加样孔的宽度一般为 1mm，图谱中 DNA 谱带宽度一般在 1mm 以上，肉眼观察 DNA 图谱，片段间相距 1mm 才能明确判定为两个基因，在低分子量区段这 1mm 相当于 50bp 差异，在高分子量区段则相当于 200bp。若两个等位基因片段长度差距小于此分辨率时，纹印图上只能观察到一条片段。因此在常规的琼脂糖凝胶上不可能分辨出多达数十个或数百个的等位基因片段。形成连续基因频率分布的另一原因是片段长度测量误差。根据同一凝胶中的标准分子量标记片段的迁移率来计算待测片段长度，迁移距离测量不可避免地会出现误差。所以单基因座 DNA 纹印的群体调查不能完全准确地确定等位基因以及基因频率分布。所测得的等位基因数是一个连续随机变量，基因频率分布也是一个变量。基因频率分布呈连续随机变量特征。

2.DNA 纹印图等位基因频率的计算方法

鉴于单基因座 DNA 纹印等位基因数是一个连续的随机变量，基因频率分布也是一个连续变量，为了将连续变量基因频率处理成不连续基因频率分布，目前应用较为广泛的是“装箱”（binning）合并的方法。即将一系列等位基因片段按从大到小的顺序分成若干组，以不连续的分组方式将连续值处理成不连续的数据群。其原则是将相邻的多态片段合并为一组，当作一个等位基因看待，计算合并后的组的频率，获得等位基因频率。目前装箱分组大致有固定箱（Fixed-bin）和滑动箱（floating bins）两种分析法。经装箱处理的单基因座探针识别

等位基因频率又称"箱频率"（bin frequency）或"固定箱频率"（fixed bin frequency）。

（1）固定箱（Fixed－bin）分析法

固定箱分析是一种合并处理办法，根据分子量标准物片段大小人为地将 DNA 图谱划分成若干个区域，每个区域称为一个"箱"，转化成一系列不连续数据。如 Lifecodes 公司提供的一套标记系统是由限制酶消化后的 λ、φ 和 T_7 病毒基因组 DNA 片段组成，共有 30 个片段，长度分别为（单位 bp）：12830，11369，10094，8453，7242，6369，5686，5220，4822，4324，3980，3675，3330，3034，2863，2693，2523，2352，2089，1925，1789，1638，1508，1353，1197，1078，964，872，773，640。30 个片段分出 31 个箱（表 10－12），每个箱的大小范围也就精确地固定下来。

表 10－12　D2S44 基因座装箱频率数据（黑人群体）

箱号	bp 范围	等位基因数	59%可信下限	点计算值	95%可信上限
1	0～639	0	.000	.000	.000
2	640～772	0	.000	·000	·000
3	773～871	4	.002	.000	.000
4	872～963	10	.008	.019	.028
5	964～1077	12	.011	.023	.045
6	1078～1196	10	.000	.019	.05
7	1197～1353	30	.035	.058	.045
8	1353～1507	27	.031	.052	.094
9	1508～1637	32	.038	.062	.087
10	1638～1688	35	.043	.068	.099
11	1789～1924	27	.031	.052	.106
12	1925～2088	36	.044	.070	.087
13	2089～2351	63	.087	.122	.108
14	2352～2522	17	.017	.033	.168
15	2523～2692	26	.029	.050	.063
16	2693～2862	23	.025	.045	.085
17	2863～3033	26	.029	.050	.078
18	3034～3329	27	.031	.052	.085
19	3330～3674	44	.057	.085	.087
20	3675～3979	11	.009	.021	.126
21	3980～4323	22	.024	.043	.048
22	4324～4821	10	.008	.019	.075
23	4822～5219	5	.003	.010	.045
24	5220～5685	4	.002	.008	.031
25	5686～6368	3	.001	.006	.028
26	6369～7241	6	.004	.012	.025
27	7242～8452	5	.008	.010	.034
28	8453～10093	0	.000	.008	.031
29	10094～11368	1	.000	.002	.000
30	11369～12829	0	.000	.000	.019
31	12830～	0	.000	.000	.000
			.000		

箱值范围确定后，凡片段长度落在同一箱内的基因，均合并为一个等位基因。合并同一箱内等位基因数则可计算该箱出现的频率。用固定箱法确定的基因频率实际上是箱的频率。因此，无论何人种群体，采用同一种分子量标准，如 Lifecodes 公司的，等位基因数都统一为 31 个。但每个箱内的片段数可能不同。固定箱分析法将连续等位基因分布简化为一张频率分布表格，它既照顾到测量误差，又得到一套不连续分布的基因频率，同时也能大致反映出基因座多态性某些特征。

装箱合并所提供的是保守的箱频率数据，但是在群体调查数据中，可能在某个箱中没有片段或仅有一条片段，致使箱频率值出现极端的低数量值，例如上述 D14S13 基因座的 1、2、28、30 和 31 箱未见有片段，29 箱仅一条片段，这 5 个箱的频率范围大约在 10^{-8} ~ 10^{-14} 之间。这种极低频率不能真正反映数据库数据，估计的等位基因频率不正确。如果出现此种情况，要求进行重新分箱，把箱内片段数少于 5 个的合并到附近的箱内，任何箱内最少有 5 条片段。将低频率的箱重新合并，使频率更加合理。这样处理之后，无论从统计学原则或实际应用原则均可以接受了。

（2）滑动箱匹配法（floating bins）

采用与计算机相连的自动搜寻光学扫描定位系统，将膜上 DNA 片段迁移的全距离人为地分成数字化位置单位，考虑到 DNA 片段长度测量误差，每一片段定位均可能与邻近片段有关。采用滑动箱确定片段位置计算基因频率之前，先应确定检测系统的片段长度测量的平均误差值。滑动箱的大小一般定在 ±1 ~ 3 个标准差范围。

采用滑动箱匹配法调查群体 VNTR 基因座频率分布的结果是一连续分布的片段分布数据，并无单个的等位基因频率。以计算机自动光学扫描的最高分辨率作为最小分格单位，经群体调查求得每个分格单位中的片段频率。其次再按检测系统的误差确定片段长度测量标准差，以 ±1 ~ 3 个标准差作为滑动箱上下限。滑动箱匹配法的优点在于充分估计测量误差因素，每一等位基因频率的确定和应用均包含了随机变量的容忍限度，即由三个可信限区间（68.3%，95.4%和 99.7%）规定的等位基因相对频率值。这样处理后等位基因频率值有所上升，同时可靠性提高了。此外，应用滑动箱法中所有的测试，其计算过程均由计算机完成，减少人为误差，利于不同实验室之间的资料交换与比较。

装箱法的目的是想把随机变量模式的 VNTR 基因座基因频率处理成不连续的数据，结果所得基因频率是一相对的数据群，从统计学原则而言，这种处理必不可少。但 VNTR 基因座具有高度多态性和高杂合度特征，往往这种装箱处理后丢失了部分遗传信息。在目前 DNA 片段分离分辨率不够高的情况下，这种处理只是一种变通的，能让法庭接受的技术性处理，其中仍然存在某种局限性。滑动箱匹配法在统计上更加可取，但需要接入一个计算机化的数据库，较复杂。固定箱匹配法只需要一个箱频率统计表，容易判读和广泛使用。

3. DNA 纹印图匹配机率的计算

无论采用固定箱或滑动箱匹配方法，将连续分布的基因频率处理为不连续数据后，即可参照共显性遗传标记系统的统计计算方法计算匹配概率。

等位基因（箱）频率 = 该基因个数/总人数 × 2；

DNA 纹印图的基因型分纯合子和杂合子，基因型频率分别为 p^2 和 $2pq$。独立遗传的基

因座图谱，累积匹配概率等于各个基因座的基因型频率的乘积。如某二个体在 D2S44 和 D7S107 基因座的基因型一致，D2S44 基因座基因型频率为 0.009636，在 D7S107 基因座为 0.005631，则在这两个基因座的累积匹配概率为 $0.009636 \times 0.005631 = 5.423 \times 10^{-5}$。

大量的 VNTR 数据表明，有些单条带的图谱实际上是杂合子，它的一个较小的等位基因因为电泳分离时已迁移出凝胶末端而未能检测出来。在这种不能判定是纯合子还是杂合子的情况下，有作者主张所有的单条带的图谱都采用 2p 来计算基因型频率，显然这是一个相对保守的值。

10.3.4 几种 DNA 纹印图介绍

国内从 1988 年开始有 DNA 纹印技术在法医学上的研究应用报道，如 D2S44（PYNH24/ HaeⅢ和 PstⅠ）（李伯龄等 1988，倪锦棠等 1993）、D2S44（Sli103/PstⅠ）、D14S1（pAC061/PstⅠ）、D14S13（pAC299/PstⅠ）、D17S79（pAC256/PstⅠ）、D18S27（pAC404/PstⅠ）及 DXYS14（pAC225/PstⅠ）（刘明俊 1989，刘明俊等 1994）、D14S1 基因座（pAW101 探针/ EcoRⅠ）（罗超权等 1992，刘超等 1994）、D14S13（pCE1.2 探针/HaeⅢ）和 D16S85（α－globin－3′HVR/ HaeⅢ）（倪锦堂等 1993）等。

1.D2S44（PYNH24 探针/ HaeⅢ或 MspⅠ）

PYNH24 特异性探针，是 Nakamura 等用人工合成的 HBV－2 寡核苷酸（GGACTGGGAGGAGTTG GGGG），从质粒 CYNH24 的人体基因组文库中分离出来长 2.0 kb 的 MspⅠ片段重组到 pUC18 质粒中，可识别 D2S44 基因座的 HaeⅢ或 MspⅠ酶解基因组 DNA 产生的限制片段长度多态性。

PYNH24 探针/ HaeⅢ酶解的 RFLP 图谱片段大小为 1.02～6.26kb，分析 146 名北京地区无关汉族个体的杂合度为 0.69，等位基因频率为 0.003～0.152，相关机率为 $3.6 \times 10^{-5} \sim 7.9 \times 10^{-2}$。50 名无关汉族个体的 PYNH24 探针/MspⅠ酶解的 RFLP 图谱片段分布在 1.5kb 至 6.5 kb 之间，也显示出高度的多态性。

2.D14S1 基因座（pAW101 探针/ EcoRⅠ）

pAW101 探针是人类 14 号染色体长臂端杜区域中的一个 5.0 kb 的片段，它与免疫球蛋白基因紧密连锁，可识别 D14S1 基因座。pAW101 探针检测 EcoRⅠ酶解基因组 DNA 的消化产物。在 106 例广州地区无关汉族个体中，检出 65 条长度各异的片段，片段长度分布在 9.4 kb～30.2 kb 之间，两个无关个体相同的概率为 0.00048。

3.D14S13（pCE1.2 探针/HaeⅢ）基因座

pCE1.2（D14S13）探针是用一段人工合成的 16mer 的寡核苷酸作为核心序列，筛选基因库，得到 Cosmid PMLJ14 后，再经 MspⅠ酶解，将其中一个 600 bp 长的 VNTR 片段克隆到 pUC18 的 AccⅠ基因座后扩增得到的，因此也称为 PMLJ14 探针。

296 名汉族无关个体的 pCE1.2 探针/HaeⅢ酶解的 D14S13 RFLP 片段长度范围为 1.08 kb～18.8 kb，杂合度为 0.94，频率分布为 0.002～0.051。等位基因频率分布峰值中可见到至少含有 66 个等位基因，且分布较为均匀，没有明显的主导基因（图 10－6）。

10.3.5 DNA 纹印图的法医学应用

1.个体识别

单基因座纹印图只涉及一个 VNTR 基因座，当两份生物检材的纹印图不一致，就可以

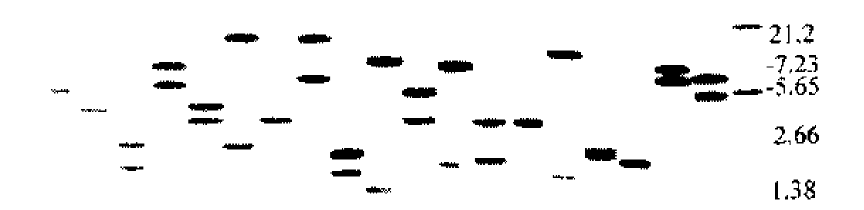


图 10-6 D14S13 基因座的谱带在 18 个无关个体中的分布

直接排除他们来源于同一个体，当两者图谱一致或匹配（match），则不排除它们来自同一个体（图 10-7），需要加测其他基因座，最后进行无关个体 DNA 纹印一致的偶合概率计算。偶合概率计算可参照共显性等位基因控制系统的偶合概率计算方法。若独立的多个单基因座纹印图（或单基因座和其它的遗传标记）一致，可按乘法原则计算总偶合率。选用 4~6 个单基因座探针，累计达到很高的个体识别能力。需要注意的是由于片段长度测量误差、部分现场物证检材 DNA 纹印存在正极漂移现象等原因，会使同一个体的现场遗留物证与新鲜血样出现差异，结果本该作同一性认定的却得出排除结论。

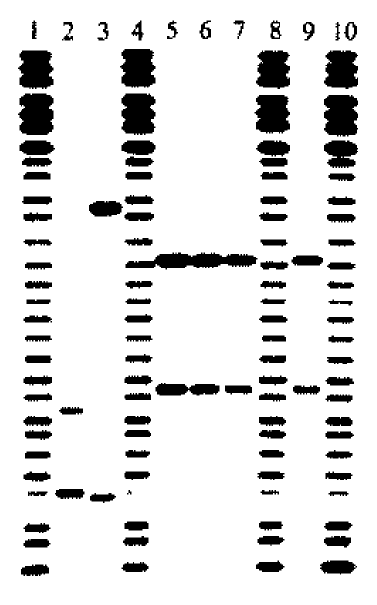


图 10-7 一起个体识别案例

编号 1、4、8、10 为分子量标准，2、3 为实验室对照 DNA，5、6、7 为从犯罪现场提取的样品，9 号为受害人的血液 DNA。

2. 亲子鉴定

单基因座 DNA 纹印的等位基因传递及基因型别类似共显性等位基因控制的血型系统，因此排除或认定亲生关系的原则，亲权指数（PI）及父子关系相对机会（RCP）计算程序与常染色体 STR 系统检验相同（详见“亲子鉴定”章）。

Balazs 计算 D2S44、D14S1、D14S13、D17S79 和 DXYS14 五个单基因座 DNA 纹印图累计非父排除率为 0.9996，几乎可以达到排除 100%非亲生父亲的准确结论。

10.3.6 应用注意问题

一个单基因座 DNA 纹印图谱中，杂合子个体只有 2 条带，纯合子只有 1 条带，根据自显影片中分子量标准物（marker）可以计算出图谱中样品片段长度。比对两图谱谱带数目与片段长度，判定两份材料是否一致，确定两份检材是否来源于同一个体。不同自显影片上的二个 DNA 纹印图谱也能进行比对，较多基因座 DNA 指纹图具有更大的优越性。

但是，在实际应用中，会遇到一些异常的 DNA 纹印图，例如有 3 条或 3 条以上的谱带或少带等情况，可能直接影响 DNA 分型结果。

1. 多带现象

(1) 混合样品　出现多带的最常见原因是多个体样品混合物。样品混合有以下几种情况：

①检验物证样品本身含有多个供体，如轮奸案中现场精斑含有几个人的精液成分，所提取 DNA 是多个人的混合样品；在几个人同时出血情况下，现场的血斑可能是几个人的混合血斑；性犯罪案中阴道分泌物与精液的混合斑。对于前两种混合样品，最后检验得到的 DNA 纹印图为混合图谱；对于阴道上皮细胞与精子混合，理论上可以采用差异裂解法，将上皮细胞 DNA 与精子 DNA 分离开来。如果检测出多个等位基因或条带，应怀疑有多个供体，一般不要人为地划分等位基因，必须对所有参与的疑犯同时进行检验，分别将嫌疑人 DNA 纹印与混合纹印作仔细比对，才能将等位基因从混合图谱中区分出来，确定罪犯的基因型。

②现场物证材料的载体本身已含有某个人的有核细胞，如受害人家中的床单上原有受害人、受害人丈夫的体液。这类检材可能会是罪犯与受害人丈夫的混合 DNA。对于这类情况，可采取空白载体对照实验，即在分析现场物证斑痕同时，剪取斑痕附近部位的干净载体进行平行分析，观察是否检出 DNA。

③在检材提取过程或在分析检验过程中由于操作不当，污染了其他 DNA，造成的混合样品。采取质量控制与质量保证体系，严格遵守操作规则，可以避免这类污染造成的混合样品。

(2) 限制性内切酶消化不完全　限制性酶切活性受到抑制，或者消化 DNA 的酶量不足，消化时间不够等原因，基因组 DNA 中一些酶切位点可能没有被切开，使 DNA 片段将比预期完全消化的长，出现额外片段。在电泳分离前用小胶电泳进行酶切程度的检验，可以消除因酶切不彻底造成的多带现象。

(3) 遗传　当 VNTR 序列内部发生点突变时，如产生一个新的 HaeⅢ酶切位点，则产生三条带。例如，正常表现为 8kb 的带，由于突变能切成二个较小谱带，如 2kb 和 6kb。另外，在某一特定基因座上，一个体可能出现三条带在遗传学上是罕见的，而且这些不常见的突变表现为孟德尔遗传、稳定传递给下一代。这种情况对于个体识别没有任何影响，在特定的案例中，这种变异成为认定个体特征的重要依据。

2. 谱带的丢失

在一些情况下，杂合子个体在 RFLP 图谱上只出现一条带，发生这种情况原因是：

(1) DNA 的降解　DNA 在体外环境如高温、高湿度、紫外线照射以及检材陈旧等引起降解；另外微生物细菌、真菌等污染，产生降解 DNA 的酶，引起 DNA 降解，失去了酶

切点；DNA 降解，产生较小片段，电泳过程中泳出凝胶，原本是杂合子的个体可能会变成纯合子，从而出现一个假排除的结论。

(2) 分析系统的分辨率　在实际分析中，RFLP 技术中的分析胶的分辨率低，不能分离两个在长度上仅差一个或几个重复单位的限制性片段，因此在图上就以单一带出现，造成该个体在该基因座表现为纯合子的假象，而事实上，该个体为杂合子。这种情况对子个体识别没有影响，而对于亲子鉴定，就要注意假排除现象。

(3) 污染　RFLP 分析依赖于酶（如 HaeⅢ、HinfⅠ等）来发挥作用。干扰酶反应的试剂将会抑制酶对 DNA 的作用。这包括各种平时经常遇到的物质，如布中的染料，尤其是斜纹布染料和一些生物物质，如血液中的血红素等，尽量在提取或纯化阶段除去这些物质。

(4) 酶切片段太小　样品 DNA 酶切片段太小，在电泳中被移出凝胶末端，只剩下较大片段，造成 DNA 谱带丢失。这种情况对个体识别与亲子鉴定影响不大。

10.3.7 单基因座 VNTR－DNA 纹印的优点及局限性

单基因座 DNA 纹印图技术与多基因座指纹图技术相比具有以下特点：

1. 所得图谱带简单明了，纯合子只有 1 条带，杂合子有 2 条带，易于理解，这在法庭辩论中是很重要的。

2. 灵敏度高，获得标准 DNA 纹印图谱的最小 DNA 量为 50ng，相当于 2μl 血液、20μl 唾液和 1μl 的精液。

3. 能够有效鉴定混合样品，纹印图中出现 3 条或 3 条以上带时，说明可能有 2 个或 2 个以上样品存在。分析样品为混合样品，如确定强奸是否为轮奸等。此外，在作大量嫌疑人筛选时，为减少工作量可以将适当数目的不同样品混合在一起加样分析，然后再对检出某一相应基因座的一组个体逐一分析，进一步确定嫌疑人。

4. 单基因座 DNA 探针在序列上具有高度的种属特异性，杂交条件为高强度杂交，人类单基因座探针不会与其他种属 DNA 杂交，细菌污染或污染有其他动物的标本，都不会影响结果的分析。

5. DNA 纹印图中 DNA 片段长度相对小（< 10kb），可用于分析部分降解 DNA 样品，获取有关的信息。

6. 检测结果重复性好，统计学评估容易理解，便于计算机管理与比对。

DNA 纹印图技术的等位基因呈连续分布，无法获得准确的等位基因频率；一个探针只检测单一一个基因座，所得信息量低，无法一次达到个体同一认定水平，需 4～6 个不同单基因座探针联合检测，费时费力，这是 DNA 纹印图技术的缺点。

10.4 DNA 指纹和 DNA 纹印操作技术

DNA 指纹和 DNA 纹印的操作程序均包括以下几个步骤：① DNA 的提取；②限制性内切酶酶切；③电泳分离；④印迹转移；⑤探针标记；⑥分子杂交；⑦ DNA 图谱的检测；⑧结果判读等（图 10－8）。

1. DNA 提取

DNA 指纹分析要求相对较大量的高分子 DNA，因此一般用有机法提取基因组 DNA。

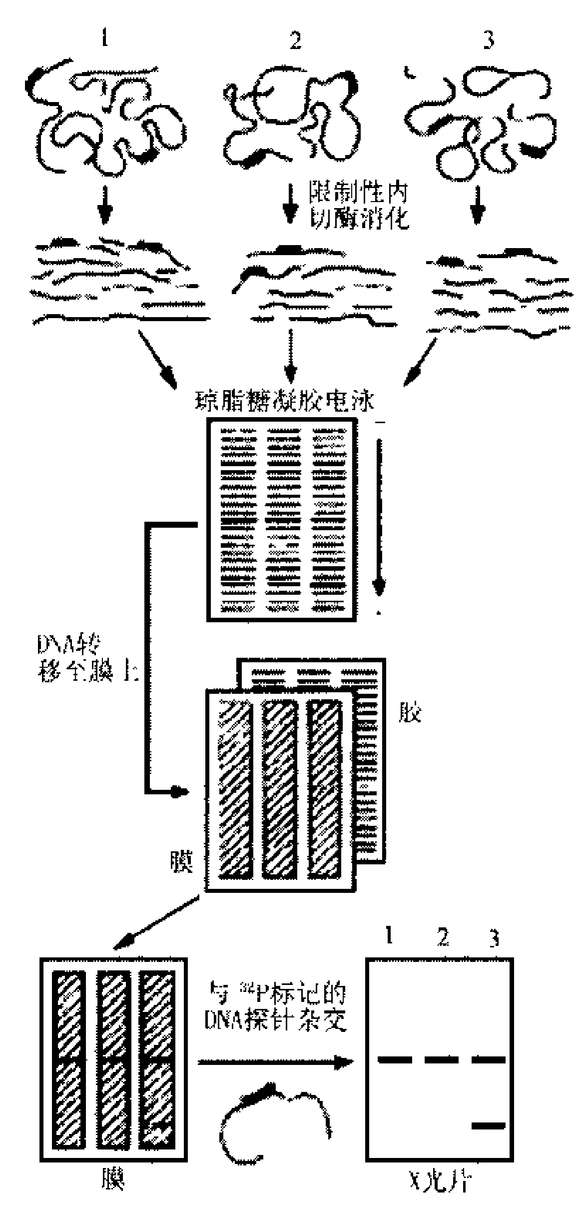


图 10－8 DNA 指纹流程图

2.DNA 浓度测定

多基因座 DNA 指纹图检测至少需要 1μg DNA，单基因座 DNA 纹印为 0.5μg DNA。如果 DNA 量太少，可能得不到完整的 RFLP 图谱。在进行 DNA 酶切前必须进行 DNA 的定量分析，以保证样品之间的一致性，同时还可以确定限制性内切酶的用量。浓度测定方法有 3 种，即分光光度法、琼脂糖凝胶电泳 EB 荧光染色法和人 DNA 特异探针斑点杂交法。具体操作方法见第五章。

3. 限制性内切酶酶切基因组 DNA

根据探针要求，选择合适的限制性内切酶。并且根据所酶切的基因组 DNA 量，确定限制性内切酶反应的体积与酶的用量。将所有反应组分混合在一起在一定温度下进行一定时间的消化，将基因组 DNA 完全消化。然后取出，加入上样缓冲液终止反应，进行电泳

分离。

计算 5～7μg 样品 DNA 的体积，计算 30μl 反应体系所需的水量，然后按下列顺序加反应组分：

10×反应缓冲液　　3μl

无菌双蒸水　　24－xμl

5～7μg 待测样品　　xμl（x＝按样品 DNA 浓度计算出的样品液 μl 数）

限制性内切酶 HinfⅠ或 HaeⅢ　3μl（30u）

混匀样品反应液，快速离心 15s，使溶液全部汇集管底部，置 37℃水浴保温 4～6h。然后取出，加入 3μl 上样缓冲液（40%甘油、0.25%溴酚蓝）终止反应，快速离心 15s，使溶液全部汇集管底部，进行电泳分离。

4. 凝胶电泳分离酶切 DNA 片段

基因组 DNA 经限制酶消化后，产生成千上万条长短不一的特异性片段。需要琼脂糖凝胶电泳将酶解片段按片段大小分离开，根据片段的迁移距离确定片段的长度。

RFLP 的靶片段长度较大，一般在 2.0～23.0kb 范围，选用 0.6%～0.8%琼脂糖凝胶电泳可获良好的分离效果，几乎都采用水平板琼脂糖凝胶的潜水式电泳。常用 Tris－醋酸（TAE）为电极缓冲液，由于 TAE 缓冲能力低，电泳槽正、负两极的缓冲液之间加一蠕动泵使之循环流动。为使大于 2kb 的片段更有效地分离，琼脂糖凝胶电泳的端电压不超过 5V/cm。电泳时间一般为 16～20h，片段小的则电泳分离时间相对短些。

（1）称取 1.2g 低电渗的粉状琼脂糖（Sigma，Ⅱ型），配制 200ml 0.6%琼脂糖凝胶，制成约 1cm 厚的平板琼脂糖凝胶。在室温下凝固 30～45min，再置 4℃下放置 10min。小心拔出梳子，去除两端的胶带或挡板，将凝胶放入电泳槽。

（2）在电泳槽两端加 1×TAE，使缓冲液液面高出胶面约 5mm。将消化好的产物与 1/10 体积的 10×载样缓冲液，混合，快速离心 15s，使溶液全部汇集管底部，加样到凝胶加样孔中。

（3）接通电流，用万用表测量并调整凝胶端电压降为 5V/cm，进行电泳。

（4）当溴酚蓝迁移到凝胶末端，切断电源，取出凝胶，进行转移。

在凝胶电泳过程中，凝胶的质量影响着 DNA 图谱的质量。薄的凝胶通常分辨率好，但载样量较少，因此凝胶一般为 1cm 厚。琼脂糖保存不好，开瓶后没有密封好，致使琼脂糖颗粒吸潮，对自显影片质量有影响，在自显影片泳道上出现背景。如果琼脂糖溶化不好，胶液中留有大量颗粒或过渡加热会降低凝胶分辨率。

另外，凝胶和电泳缓冲液中含有染料 EB，DNA 被 EB 结合滞后迁移率，当缓冲液不能循环时，较小片段由于失去 EB 迁移更快。

5. 印迹转移

凝胶电泳后，DNA 片段按片段长度排列在凝胶中，由于琼脂糖凝胶的机械强度较差，易破碎，而且分子杂交后背景高，需要将凝胶中的 DNA 转移到一种韧性好背景低的固相支持物上。常采用真空转移方法将凝胶中的 DNA 转移到尼龙膜上。而且，酶切片段较大，为了提高大片段的转移效率，在碱变性前进行 0.25mol/L HCl 脱嘌呤处理，打断大分子 DNA 以提高大片段的转移效率。DNA 片段从凝胶中原位转移固定在尼龙膜转移过程中，

各 DNA 片段的相对位置保持不变，该膜永久性地重现出 DNA 片段在凝胶中的位置，便于下一步的杂交反应。

（1）在真空转移仪的真空密封膜上剪一比凝胶稍小（0.5cm 左右）的窗口。

（2）裁剪一比窗口稍小的尼龙膜或硝酸纤维素膜，用铅笔作上标记，用去离子水充分湿润，置于上述密封膜的窗口内，排除两者之间的气泡。

（3）取出电泳后带凝胶的制胶盒，一手挡住凝胶的一端稍微倾斜，将凝胶置于尼龙膜或硝酸纤维素膜上。排除气泡。

（4）接通真空泵，真空压约为 20cm H_2O 柱。

（5）用 0.25mol/L HCl 酸变性液覆盖住凝胶，抽气 3min，真空压为 40cm H_2O 柱。

（6）抬高真空转移仪的一端，吸去酸变性液，改用碱变性液，抽气 4min，真空压为 40cm H_2O 柱。

（7）吸去碱变性液，换用中和液，同真空压下继续抽气 3min。

（8）吸去溶液，用转移液覆盖凝胶，60～90cm H_2O 柱真空压下抽气 40min。在酸、碱处理及转移过程中，如果溶液减少了，需随时添加转移液。

需要注意的是，真空转移时间不宜过长。如果转移时间过长，DNA 片段有可能通过尼龙膜，降低灵敏度。转移时凝胶和膜之间如有气泡或滤纸，影响转移效果，在自显影片上这些部位会出现花斑。

（9）碱固定　转移完毕，取出尼龙膜在 2×SSC 中漂洗 1min。随后将膜结合 DNA 面向上，平铺于用 0.4N NaOH 浸透的三层滤纸上，去除滤纸与膜之间气泡，固定 2min，再于适量 5×SSC 中漂洗 1min，取出，用滤纸吸干，备用或直接用于杂交。

6. 探针标记、杂交与检测

探针标记与杂交已在前面章中详细介绍了，DNA 指纹中的探针多为小卫星探针，随着合成技术与标记技术的提高，目前大多根据小卫星探针的核心序列合成寡核苷酸探针，采用辣根过氧化物酶或碱性磷酸酶等直接标记探针，进行非同位素探针杂交检测。

（1）取 5μl（500ng）于 0.5ml Eppendorf 管中，用无菌去离子水稀释至 10ng/μl。

（2）沸水中煮沸变性 5min，取出迅速插入冰浴中 5min。

（3）加入 45μl HRP 标记试剂，充分混匀，稍离心。

（4）加入 50μl 戊二醛充分混匀，置 37℃水浴保温 15min 后，可以直接用于杂交。如果不当时使用，可加入甘油使其终浓度为 50%，于 －20℃保存。辣根过氧化物酶标记的探针不需要进行纯化，可直接用于杂交。

7. 杂交

（1）试剂

①预杂交液　ECL 试剂盒附带的杂交液 40ml，加入 1.17g NaCl，使终浓度为 0.5mol/L，再缓缓在搅拌下加 2g 酪蛋白，60℃加热溶解。

②杂交液　上述预杂交液中加入 500ng HRP 标记的变性探针 DNA。

（2）杂交

将已转移的滤膜卷成圆桶状，放入杂交管中，加预杂交液 40ml，去除管壁与滤膜之间的所有气泡，拧好盖，放入 41.5℃的杂交箱中预杂交 0.5～1h。预杂交结束后，加入

150μl HRP 标记的探针 DNA，继续在 41.5℃杂交过夜。

辣根过氧化物酶对温度较敏感，超过 42℃易失活，在杂交过程中需注意控制温度。

(3) 后处理

杂交结束后，倒去杂交液，取出膜，放入洗膜液 I （6mol/L 尿素，0.5×SSC 和 0.5% SDS）溶液中，室温洗涤 25min，重复一次；再转到 2×SSC 溶液中，室温洗涤 5min。取出膜放在滤纸上，然后进行检测。

8. 化学发光法检测

(1) 将杂交、洗涤后的尼龙膜用滤纸吸干，平放一干净玻板上，DNA 面朝上，取试剂盒中的试剂 1（鲁米诺）和 2（过硼酸盐）等体积混匀，覆盖于尼龙膜上，反应 1min，用滤纸吸去反应液，用保鲜膜将膜包好，置于暗盒中再压上感光胶片，关上，室温下曝光。

(2) 20min 后取出胶片，显影、定影，得到 DNA 图谱。

需要注意的是，在检测自显影时，增感屏被污染，不论是否放杂交膜，自显影片上出现斑点，如果膜在放人杂交液前被放射性同位素标记或酶标记的探针溶液污染，会出现斑点。如果洗膜时膜粘在一起，洗膜没洗好，或在杂交过程中膜被镊子损坏，或是膜上长了真菌，或是细菌碱性磷酸酶活性会引起整个照片高背景，使自显影照片不清晰，背景高。

第十一章　小卫星 VNTR

11.1 VNTR 概述

基因组 DNA 中存在一类串联重复序列，其串联重复单位（核心序列）数目在人群中存在较大差异，具有高度多态性，称此重复为可变串联重复序列（VNTR)，根据重复单位长度，分为二类：重复单位长度为 10～70bp，称为小卫星，长度为 1～7bp 的，称为微卫星，或叫 STR。

根据 VNTR 基因座侧翼序列，设计、合成与侧翼互补的寡核苷酸作为 PCR 引物，应用 PCR 技术扩增包含 VNTR 基因座侧翼和重复单位序列，凝胶电泳分离片段长度不同的扩增产物，根据扩增片段长度差异确定其基因型。因此，称这种用 PCR 技术检测 VNTR 基因座多态性的分析为扩增片段长度多态性（amplified fragment length polymorphism，AMP－FLP）分析，它充分发挥了二个优势：一是基因组 VNTR 基因座具有高度多态性优势，二是 PCR 快速、高效扩增的优势。AMP－FLP 分析是 20 世纪 90 年代法医 DNA 分析技术一项突破性进展，使 DNA 分型实现了高效、快速、灵敏的目标，被称为第二代 DNA 分型技术。

AMP－FLP 图谱简单，纯合子个体只有一条带，含有二个长度相同的等位基因；杂合子个体有二条带，含有二个长度不同的等位基因。法医学鉴定应用 AMP－FLP 有以下几个特点：

1. 高灵敏度。PCR 技术理论上可以扩增出单个细胞基因组 DNA，在日常物证鉴定中，灵敏度可以达到 ng～pg 级。

2. 适用于陈旧、降解、腐败检材。小卫星 VNTR 基因座多态片段长度多在 1kb 以下，微卫星座位的靶片段一般不超过 400bp，因此模板 DNA 只要含有引物结合序列及目的重复序列区就可以得到扩增产物。

3. 可以进行多个基因座复合扩增检测。PCR 检测的模板 DNA 用量少，可以进行多次或多个基因座检验，也可以将扩增条件近似的基因座安排在同一个反应体系中进行复合扩增，增加单次检测的信息量。

4. 具有种属特异性。在引物设计上选择人类特异的序列，其它物种 DNA 一般不会得到扩增。

5. 能够进行混合检材的个人识别。对于单一个体，PCR 扩增产物只有一条或二条谱带。如果有二个或二个以上不同个体来源 DNA 混合在一起，扩增图谱中会出现三条以上的带。由于 PCR 反应的高灵敏度，即使二个个体的 DNA 按 1:10 混合，也能分辨出来。

6. 可获得不连续的等位基因频率。选用的 VNTR 座位的等位基因片段长度在 1kb 以下，等位基因片段间按重复单位数成等差数列，电泳分离后等位基因判定明确。等位基因分布呈不连续性，群体等位基因频率可以按显性基因统计方法计算。

7. 实验周期短，设备条件和操作相对简单。

迄今，已有许多小卫星 VNTR 基因座应用于法医 DNA 分析中，如 D1S80，ApoB，pYNZ22 等基因座（表 11－1）。本章将详细介绍扩增片段长度多态性分析小卫星 VNTR 基因座。

表 11－1 小卫星 VNTR 位点与基因组内的定位

小卫星 VNTR 基因座名称	基因组内的定位
D1S80 （pMCT118）	1P
ApoB	2 号染色体
D17S30 （pYNZ22）	17 号染色体
D1S118 基因座 （p33.4）	1 号染色体
D1S111 （p 33.6）	1 q21－1q31
COL2A1 人类胶原蛋白Ⅱ型基因（collagen Ⅱ gene）	12q14.3
IL－6 白细胞介素 6 基因（Interleukin－ 6 gene）	7p15－p21
D4S95	4 号染色体
DXS52	Xq28

11.2 小卫星 VNTR 基因座分型基本技术

11.2.1 试剂

1.DNA 提取试剂：提取试剂见 DNA 提取章节。

2.引物：寡核苷酸引物长度为 18～20 个碱基，为人工合成，合成后的引物一般制成冷冻干粉，使用前先加超纯水溶解，待引物溶解后测定 260nm 处的消光吸收值，按公式计算引物稀释液浓度（μmol/L）$= \frac{\text{O.D 值} \times 40 \times \text{稀释倍数}}{\text{引物碱基数} \times 323}$

上式中的 40 为单链 DNA 在 260nm 波长处消光值为 1OD 时的浓度，323 为一个碱基的平均分子量。

按 PCR 反应的 10 倍工作浓度稀释引物，分装，－20℃贮存备用。避免反复冻融，否则引物易降解，PCR 扩增效率下降。

3.PCR 缓冲液：标准 1×PCR 缓冲液组分为：10mmol/L Tris－HCl（pH8.3），50mmol/L KCl，1.5mmol/L $MgCl_2$，0.001% 白明胶。一般购买 Taq DNA 聚合酶时带有 10× PCR 缓冲液，缓冲液成分可能有所不同。

4.dNTPs：dNTPs（包含 dATP、dTTP、dCTP、dGTP）贮存液浓度为 10mmol/L 或 2.5mmol/L，工作液浓度一般为 200μmol/L

5.Taq DNA：聚合酶贮存液浓度为 5u/μl

6.特异基因座等位基因分型标准物（Ladder）：通过人群调查筛选出所有等位基因片段，经序列分析确认，大量扩增，然后混合作为等位基因分型标准物，也可用公司制品。

7.琼脂糖

8.30%丙烯酰胺溶液

9.2%亚甲基双丙烯酰胺溶液

10.10%过硫酸胺溶液，TEMED

11.5 ×载样缓冲液（pH8.0）：50%甘油，0.4%溴酚蓝，0.1%二甲苯菁蓝，1mmol/L EDTA

12.10 × TBE 溶液（pH 8.0）：0.9mol/L Tris，0.6mol/L 硼酸，2mmol/L EDTA

13. 银染固定液：10%乙醇，5%乙酸

14. 银染染色溶液：12mmol/L $AgNO_3$

15. 银染显色溶液：0.25mol/L Na_2CO_3，0.04%甲醛（1ml 40%甲醛到 1 000ml 溶液）

16.10%乙酸溶液

11.2.2 模板 DNA 制备与浓度测定

模板 DNA 制备及 DNA 浓度测定参见 DNA 提取与定量部分。对于小卫星 VNTR 基因座采用酚-氯仿法提取 DNA 较好。如果用 Chelex-100 方法提取，某些基因座扩增可能不太理想。

11.2.3 PCR 扩增反应

标准 PCR 反应体系一般为 50μl 或 25μl，在一个反应管中将 PCR 反应各组分混合后置热循环仪中，设定热循环参数进行扩增。

1. 根据扩增仪型号选择 0.5ml 或 0.2ml 离心管，标记好每一个离心管。

2. 模板 DNA 用量 1～100ng，对于有机溶剂提取、酒精沉淀的模板 DNA，先稀释成 1～100ng/μl，后取 1μl 用于 PCR 反应；如果为 Chelex-100 提取液，一般取 5～7μl 用于 PCR 扩增。

3. 确定反应管数，然后按表 11-2 各反应组分，计算 10× 缓冲液，dNTPs，引物，及 Taq DNA 聚合酶与灭菌超纯水等总量，表 11-2 是以 D1S80 基因座为例设定的浓度与用量，其他基因座可参照设定。

表 11-2　D1S80 基因座的 50μl 扩增体系各组分含量

PCR 反应液组分	贮液浓度	需加的量
1×缓冲液	10× 缓冲液	5μl
200μmol/L dNTPs	2.5mmol/L	4μl
0.2μmol/L 引物 1	2μmol/L	5μl
0.2μmol/L 引物 2	2μmol/L	5μl
模板 DNA	1～100ng	xμl
Taq DNA 聚合酶 1.25u	5u/μl	0.25μl
灭菌超纯水		至 50 μl

4. 按下列顺序加样：先加灭菌超纯水、缓冲液、dNTPs，引物及 Taq 酶，然后旋涡振荡数秒混合均匀，快速离心数秒。

5. 按每份（50-x）μl 分装到每一离心管，x 为模板 DNA 用量。

6. 最后加入模板 DNA，混匀，加 1～2 滴无菌的石蜡油覆盖 PCR 反应液。如果扩增仪配有热盖，使用薄壁离心管则无须加石蜡油。

7. 最后快速离心数秒，放入扩增仪内，设定 PCR 热循环参数，进行扩增。

标准的循环参数为 95℃变性 5min 后，94℃变性 1min，60℃退火 1min，72℃延伸 3min，循环 35 次后退出，72℃延伸 7min，最后 4℃保温。

11.2.4 PCR 扩增产物的电泳分离与检测

不同片段长度的 PCR 产物经电泳分离后按片段分子大小彼此分离，通过谱带显现，

可以确定等位基因基因型。PCR 扩增产物电泳分离方法主要有：琼脂糖凝胶电泳和聚丙烯酰胺凝胶电泳。可根据扩增产物片段长度范围分别选用。

1. 琼脂糖凝胶电泳与 EB 染色法

（1）小卫星 VNTR 基因座扩增片段长度一般在 300～1 000bp 之间，配制 1%～2%琼脂糖凝胶，凝胶缓冲液为 1×TBE，凝胶长度为 14cm 以上，凝胶配制方法参见第五章。

（2）取 10～15μl PCR 扩增产物与等量载样缓冲液混合加样，1×TBE 溶液为电极缓冲液，在 100V 或 200V 恒电压条件下电泳，电泳时间根据检测基因座片段长度确定，一般可根据载样缓冲液中染料溴酚蓝与二甲苯青蓝的位置判断。

（3）采用 EB 染色，紫外灯下观察电泳分离的 DNA 片段，照相记录结果。

EB 溶液直接加在琼脂糖凝胶中，这样在电泳过程中能随时观察电泳分离情况，但容易污染环境；也可以在电泳结束后，将凝胶浸在 EB 溶液中，然后取出观察。琼脂糖凝胶电泳与 EB 染色请参见第五章。也可以用 SYBR Green 代替 EB。

2. 垂直非变性连续聚丙烯酰胺凝胶电泳分离

利用 T5%的变性或非变性聚丙烯酰胺凝胶垂直电泳，银染显色方法检测扩增产物。在 20 世纪 90 年代初期，小卫星基因座的扩增产物大都采用微型聚丙烯酰胺凝胶（10×8cm 左右）电泳，具有快速、灵敏、操作简便等特点，但分辨率低，目前都采用长距离（14～20cm）的变性聚丙烯酰胺凝胶电泳。具体方法同 STR 基因座。

结果的判读是以已知重复单位数的等位基因分型标准物（allele ladder）做为参照进行的，以等位基因重复数命名等位基因。如样品的谱带位置与邻近的等位基因分型标准物中等位基因 6 一致，则样品该谱带为等位基因 6。

11.3 常用小卫星 VNTR 基因座

11.3.1 D1S80 基因座

D1S80 基因座（pMCT118）定位于染色体 1p，是法医 DNA 分析中研究应用较早的一个 VNTR 基因座。核心序列为 GNNGACCACCGGNAAG。经双脱氧链终止法测得 D1S80 基因座的核苷酸序列如下：

AGGCT CCCAC CTCAG GGTGC TCGCT TGGAG CAAAC AGGCC AGAGG GGAGC GAACA CAGTG? TGAG GGGCCGATGGACGCT TCCAC TGGCC TCCCC TCATC GACCG GCCCC TCACG GTGCC AAGGAAACAG CCCCA CCCATG AGGCG CTGAG AGAAACTGGCCTCCA AACAC TGCCC GCCGT CCACG GCCGG CCCGTCCTGCGTGTG AATGA CCCAG GAGCG TAT? CCCCAC GCGCC A? ACT GCGAT CAGAT AAGCG

GCAGC CCA－A GG－AA C

ACAGA CCACC GGCAA G

GAGGA CCACC GGAAA G

GAAGA CCACC GCAAA G

GAAGA CCACC GGAAA G

GAAGA CCACC GGCAA G

GAGGA CCACC GGAAA G

GAAGA CCACC GGCAA G

GAGGA CCACC AGGAA G

GAGGA - - - - - - - - - - -

- - - - - - - - - - - - - - - GGAA G

GAGGA CCACT GGCAA G

GAAGA CCACC GGCAAA G

CCTGC AAGGGG GCACC TGCAT CTCCA ACAAG ACAAA ATAAA CAAGC CAGAG AGGGC TTGTG ACCAG TGTGC CATTT GTCAC。

Kasai（1989 年）首先报道 PCR 扩增 D1S80 基因座获得成功，证实该基因座重复单位长度为 16bp。目前已发现 36 个等位基因（等位基因 13～45，50），等位基因片段长度范围为 260～900bp。其中等位基因 45 和 50 是 1999 年才报道发现的。

国内李伯龄等（1991）首先报道了对该基因座的应用研究，在 98 名无关汉族个体中查出 22 个等位基因。迄今在中国汉族人群中已发现 29 个等位基因，为等位基因 14～45。基因座杂合度大于 0.84，DP 值为 0.96，具有高度的多态性。不同群体等位基因分布频率见表 11－3。

表 11－3　D1S80 基因座群体等位基因分布频率

等位基因	群体分布频率						
	中国汉族（沈阳）	中国汉族（浙江）	中国汉族（台湾）	中国白族（云南）	中国壮族（广西）	白人	黑人
14	0.0033	0.0048	0.005			0.0044	
15				0.0086	0.002		
16	0.0246	0.0048	0.010	0.0043	0.009		0.002
17	0.0083			0.0129	0.003		0.039
18	0.2110	0.1154	0.129	0.2414	0.198	0.2477	0.069
19	0.0299	0.0385	0.014	0.0129	0.018	0.0044	
20	0.0050			0.0086	0.011	0.0132	0.022
21	0.0249	0.0241	0.038	0.0172	0.030	0.0310	0.163
22	0.0150	0.0096	0.014	0.0259	0.023	0.0663	0.071
23	0.0216	0.0096	0.10	0.0216	0.031	0.0177	0.035
24	0.1977	0.2644	0.276	0.1983	0.225	0.3496	0.200
25	0.0332	0.0288	0.050	0.0690	0.036	0.0310	0.051
26	0.0033		0.005	0.0905	0.045	0.0177	0.014
27	0.0266	0.0337	0.062	0.0345	0.063	0.0044	0.004
28	0.0847	0.0769	0.067	0.0690	0.060	0.0354	0.133
29	0.0465	0.0337	0.024	0.1164	0.035	0.0398	0.045
30	0.1063	0.1683	0.148	0.0345	0.115	0.0089	0.014
31	0.1046	0.1346	0.114	0.0086	0.052	0.0663	0.055
32	0.0166	0.0192	0.014	0.0216	0.023	0.0089	0.004
33	0.0166	0.0048	0.014	0.0126	0.023	0.089	0.0067
34	0.0066					0.0177	0.004
35	0.0066	0.0048		0.0043	0.002		0.004

续表

36	0.0100		0.010		0.008	0.0132	
37					0.005	0.0132	
38	0.0017						0.002
39	0.0017				0.003		
40						0.0044	
41	0.0083	0.0096				0.0044	
41a							
42		0.0096			0.003		
43							
44	0.0017						
45		0.0048					
杂合度	0.861	0.8625	0.905	0.8543	0.834	0.788	
DP		0.9501			0.958		

对 D1S80 基因座一些等位基因序列测定分析发现，D1S80 基因座不仅有重复单位数目的差异性，而且在一些重复区域有核苷酸的取代，出现一些限制性内切酶的切点，具有碱基序列多态性，可以采用 PCR－RFLP 或 PCR－MVR 方法测定，提高个人识别力。如在 D1S80 基因座 5′区域有 HinfⅠ$^{+}$/HinfⅠ$^{-}$和 Tsp509Ⅰ$^{+}$/Tsp509Ⅰ$^{-}$等二等位基因多态性，增加了基因座的遗传多态性。图 11－1 中 A、B、C 样品在 AMP－FLP 图谱中表现为相同基因型，进一步用内切酶酶切 PCR 产物后，三个样品的电泳图谱各不相同，表现出差异性。

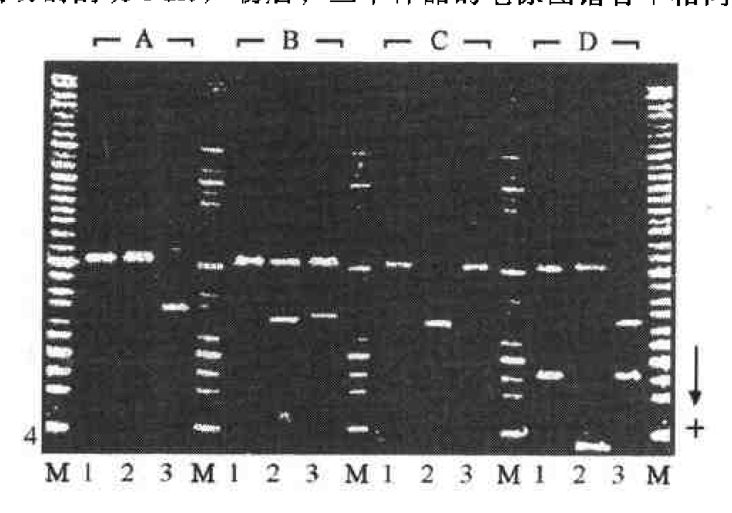


图 11－1 D1S80 基因座的 PCR－RFLP 分析

1：未经处理样品；2：HinfⅠ酶切；3：Tsp509Ⅰ酶切；M：D1S80 等位基因分型标准物（Ladder）

A：等位基因 24T－24T；B：24G－24T；C：24G－24G；D：17G－24T

Antonio Alonso 报道正向链引物 5′端第 61 核苷酸有 C→T 的转换，产生 HinfⅠ的酶切位点。在西班牙人群中 HinfⅠ$^{+}$和 HinfⅠ$^{-}$频率分别为 0.49 和 0.51，杂合度为 0.50，并发现等位基因 18 与 HinfⅠ$^{+}$相关，等位基因 24 与 HinfⅠ$^{-}$相关，其它等位基因的 HinfⅠ$^{+/-}$表现为随机性（图 11－2）。吴微微等人（1998）报道在中国汉族群体中 HinfⅠ$^{+}$和 HinfⅠ$^{-}$频率分别为 0.70 和 0.30。但是 G. Watanabe（1997）测序证明是正向链引物 5′端第 58 位核

苷酸发生 G→T 的颠换（图 11－3），而不是第 61 位核苷酸。等位基因 16 的第 58 个核苷酸为 G，有 Hinf Ⅰ 的切点，而等位基因 24 的则为 T。另外，第 58 位 G→T 转换产生 Tsp509I 酶切多态性。

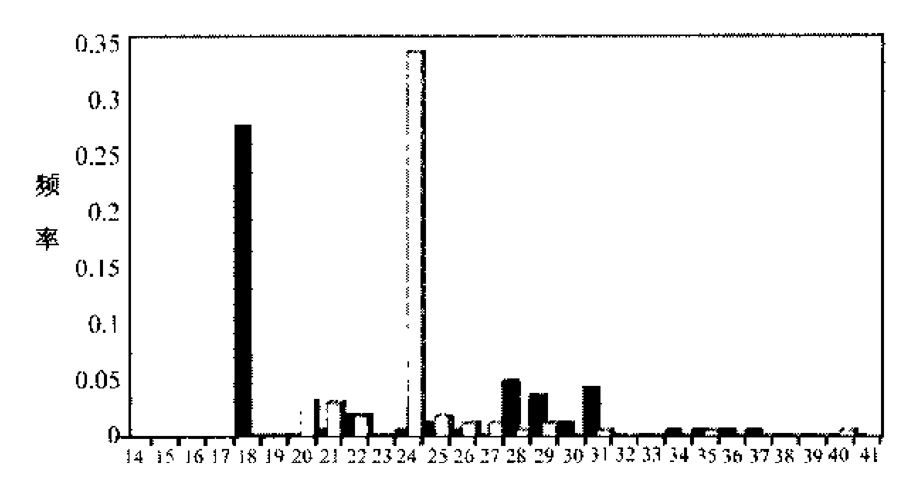


图 11－2　D1S80 基因座 HinfⅠRFLP/VNTR 单倍体型的频率（82 个西班牙人）

（■）Hinf Ⅰ＋/VNTR 单体型，（□）Hinf Ⅰ－/VNTR 单体型

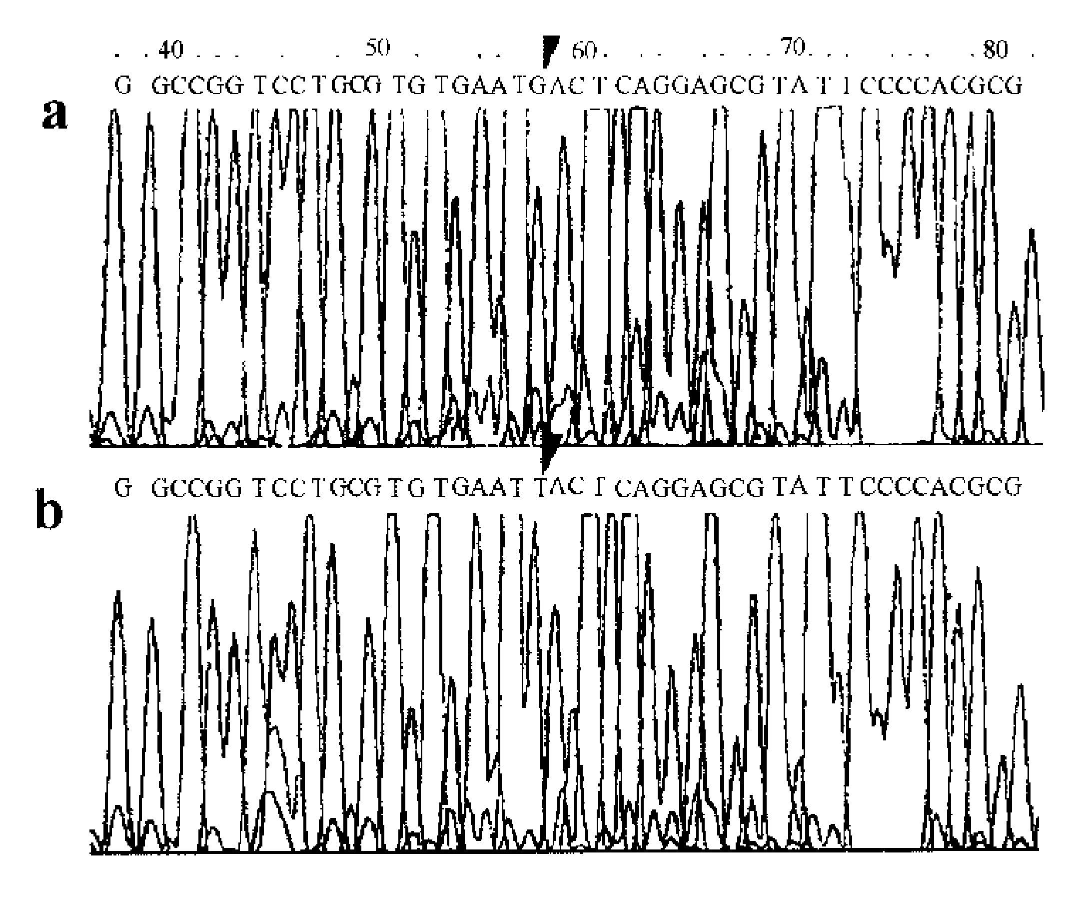


图 11－3　D1S80 基因座侧翼序列分析结果（正链引物 5′端的第 38－82 个核苷酸）

a：D1S80 基因座的等位基因 16　b：等位基因 24；图中箭头所指为突变点

从日本人群体单倍体型数据分析，除等位基因 24 属于 T 型外，大多数等位基因属于 G 型，G 型是 D1S80 基因座原始的 VNTR 等位基因，进化中发生了 G－T 的转换，产生了等位基因 24 的二态性特征。Albarran 等人认为等位基因 18 和 24 可能是 D1S80 基因座原始的 VNTR 等位基因，含有 G 型的等位基因来自于等位基因 18，含有 T 型的等位基因则来自于等位基因 24。

除了第 58 位核苷酸 G—T 颠换，还有其它位置的碱基序列变化。Harashim（1997）在分析 180 个日本人时发现在 18 个等位基因 27 中，有 5 个等位基因为等位基因 27 的变异体，在 6% 非变性凝胶中电泳位于等位基因 26－27 之间，与等位基因 27 很接近，它们碱基数均为 578bp，与标准等位基因 27 重复单位组成相同，但序列不同，碱基变化只发生在重复单位内，侧翼序列与标准等位基因没有区别，可分成 V1、V2 和 V3 三个亚类（图 11－4，5）。与标准等位基因比较，V1、V2、V3 碱基取代分别发生在第 18、12、8 个重复位置（图 11－6）。V1 中有 12 个不同的重复子，V2 有 7 个，V3 有 6 个不同重复子，从这个观察结果预测 D1S80 基因座其他等位基因的变异体、如等位基因 17，22，23，24，25 或 31 也可能由相同重复单位组成，但重复单位排序不同。用 EcoRⅡ酶切所有等位基因 27，发现标准等位基因 27 与变异体的酶切图谱不同（表 11－4），317bp、269bp、253bp 片段在 PAG 图上迁移率不同（图 11－7），可以区分变异体。V3 和标准等位基因 27 比较，有 269bp，没有 48bp 的带，区别于变异体。用 MspⅠ酶切，V1 产生 9 个片段，V2 为 8 个，V3 为 7 个片段，192bp 只产生于标准体中，因 V3 没有 64bp 和 48bp 片段可以与 V1、V2 区别开，V2 因缺乏 32bp 片段区别于 V1。

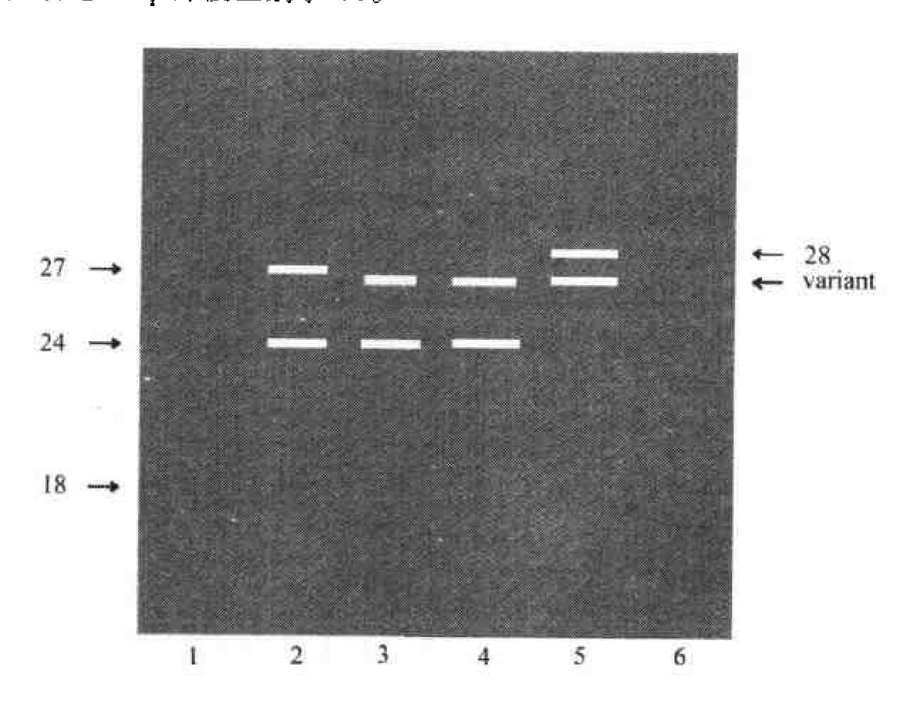


图 11－4 D1S80 基因座等位基因 27 和变异体的聚丙烯酰胺凝胶（6%）电泳图

2：标准等位基因 27/24；3：变异体 1/24；4：变异体 2/24；5：变异体 3/28

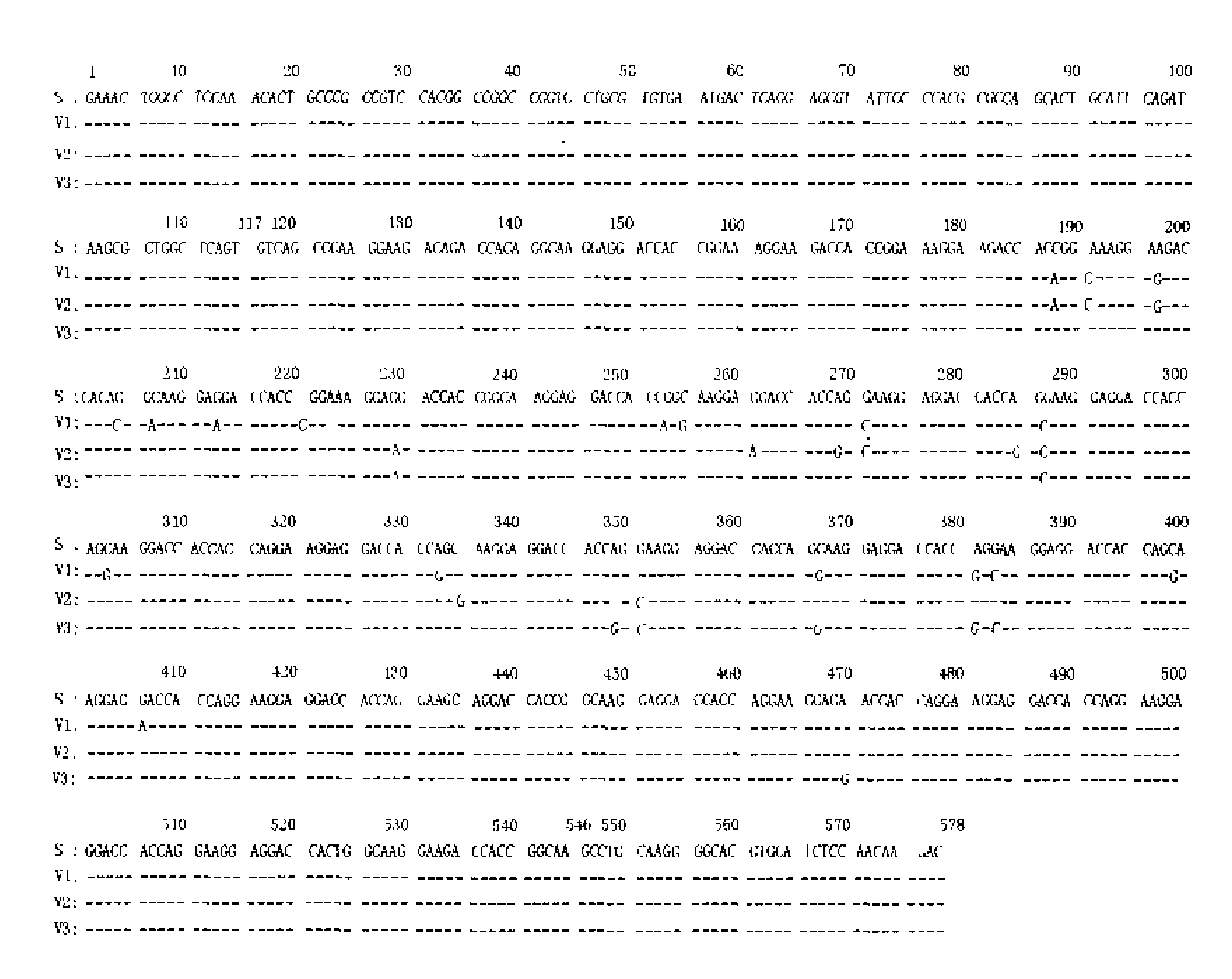


图 11－5 D1S80 基因座等位基因 27 和三种变异体的序列

S：标准等位基因 27；V1－V3：变异体

Rep. No.	1	2	3	4	5	6	7	8	9	10	11	12	13	11
S:	1	2	3	4	4	5	3	8	8	9	9	10	9	10
V1:	1	2	3	4	5	3	7	8	9	10	10	9	9	8
V2:	1	2	3	4	5	6	3	7	8	7	8	10	9	9
V3:	1	2	3	4	4	5	3	7	8	9	10	10	9	10

Rep. No.	15	16	17	18	19	20	21	22	23	24	25	26	27
S:	9	10	9	10	9	9	8	9	11	9	9	12	7
V1:	9	9	8	9	11	9	8	9	11	9	9	12	7
V2:	8	10	9	10	9	9	8	9	11	9	9	12	7
V3:	8	9	8	10	9	9	8	9	9	9	9	12	7

1：TCAGC CCAAG GAAG
2：ACAGA CCACA GGCAA G
3：GAGGA CCACC GGAAA G
4：GAAGA CCACC GGAAA G
5：GAAGA CCACA GGCAA G
6：GAGGA CCACA GGCAA G
7：GAAGA CCACC GGCAA G
8：GAGGA CCACC GGCAA G
9：GAGGA CCACC AGGAA G
10：GAGGA CCACC AGCAA G
11：GAGAA CCACC AGGAA G
12：GAGGA CCACT GGCAA G

图 11－6 D1S80 基因座标准等位基因 27 和变异体的重复单位序列

S：标准等位基因 27；V1－V3：变异体

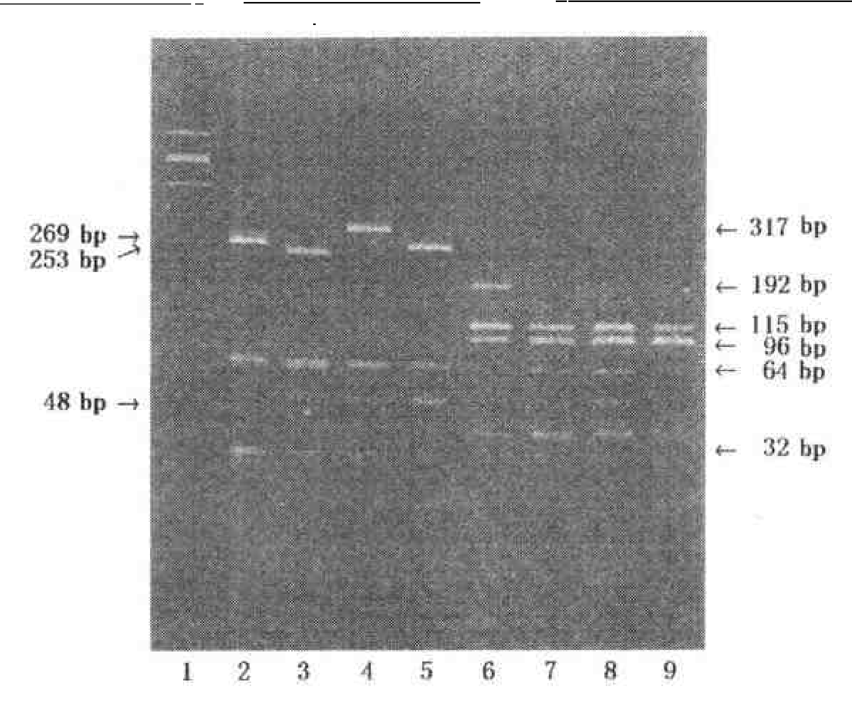


图 11-7 标准等位基因 27 和变异体 EcoR Ⅱ 或 Msp Ⅰ 酶切图谱

1：PBR322-HaeⅢ；2-5：EcoRⅡ处理；6-9：MspⅠ处理

表 11-4 扩增片段限制酶切片段大小

片段大小 (bp)	EcoRⅡ				MspⅠ			
	S	V1	V2	V3	S	V1	V2	V3
317			+					
269	+			+				
253		+						
192					+			
115					+	+	+	+
96					+	+	+	+
70	+	+	+	+				
64						+	+	
48		+	+	+		+	+	
39					+	+	+	+
38					+	+		+
32	+	+	+	+	+	+	+	+
16	+	+	+	+	+	+	+	+
4					+	+	+	+

（1）引物序列　5′-GAA ACT GGC CTC CAA ACA CTG CCC GCC G-3′

5′-GTC TTG TTG GAG ATG CAC GTG CCC CTT GC-3′

（2）PCR 扩增反应　50μl 反应体系，含有 0.4～5ng 模板 DNA，10mmol/L Tris-HCl (pH8.3)，50mmol/L KCl，1.5mmol/L $MgCl_2$，0.001% 白明胶，200μmol/L dNTPs，0.2μmol/L 的每种引物，1.25u Taq 聚合酶。

（3）PCR 热循环参数　95℃变性 2min，然后按 95℃45s，67℃ 1min，70℃ 3min，扩增

循环 35 次，最后 72℃ 7min，退出循环。

或 95 ℃ 1min，65℃ 1min，72℃ 1min，30 个循环，最后 72℃ 7min。

11.3.2 APOB 基因座

ApoB（apolipoprotein B）基因座位于人类染色体 2p24 - p23（GDB：4230792），是低密度脂蛋白受体的配体，以 ApoB - 100、ApoB - 48 两种形式存在。ApoB VNTR 位于载脂蛋白 B 基因 3′末端码区的 0.5kb 处，这是一个富含 A - T 碱基的小卫星区域，含有 30 个碱基对，由两种结构的亚单位串联重复序列构成。这两个亚单位分别是 X 和 Y，序列为 ATAATTAAAATATTT 。最早由 Eric Boeiwinkle（1989）用 PCR 分析 3′ApoB 多态性，目前已发现有 24 个不同的等位基因，片段长度为 570 ~ 900bp 或 580 ~ 910bp（J.Schnee - Griee，1991），等位基因频率见表 11 - 5。国内自从李伯龄等（1992）报道在 100 例北京地区汉族个体查出 11 个等位基因以来，至今已发现 16 种等位基因（等位基因 27 ~ 50），43 种基因型，基因频率为 0.0025 ~ 0.3750，最常见的等位基因为 35（频率 0.3750）和 33（频率 0.1825），频率最低的等位基因为 27（0.0025）。最常见的基因型为 35 - 35、31 - 35 和 33 - 35。杂合度为 0.79，个人识别机率（DP）为 0.95。然而，其他人群中等位基因 37 最为常见。

表 11 - 5　APOB 基因座等位基因频率

等位基因	频率				
	Renges（亚洲人）	德国人	法国人	西班牙人	美国黑人
15	0.0005				
17	0.0009				
22				0.0027	0.031
25					0.005
26					0.005
28					0.005
29	0.009		0.004	0.0055	0.036
30			0.000	0.0055	0.020
31	0.131	0.093	0.065	0.0765	0.046
32	0.005		0.000		0.010
33	0.061	0.040	0.054	0.0574	0.097
34		0.002	0.000		0.005
35	0.364	0.254	0.217	0.2322	0.168
36	0.005	0.006		0.0027	0.005
37	0.248	0.370	0.404	0.3852	0.199
38	0.009		0.000		0.005
39	0.042	0.037	0.052	0.0301	0.092
41		0.0123	0.023	0.0191	0.097
43		0.002	0.002	0.0191	0.071
45		0.007	0.004	0.0574	0.046
47	0.042	0.072	0.071	0.0956	0.026
49	0.061	0.088	0.077	0.0109	0.026
51	0.009	0.013	0.027		0.000
53		0.006			0.005
H				0.7764	
DP				0.9238	

注：中国汉族群体数据虽有许多文章报道，但由于没有标准的等位基因参照物，因此这些数据频率未汇入此表。

（1）引物序列 5′－ATG GAA ACG GAG AAA TTA GC－3′

5′－CCT CTC ACT TGG CAA ATA C－3′

或 5′－TGG AAA CCG AGA AAT TAT GGA GC－3′

5′－CCT TCT CAC TTG GCA AAT ACA ATT－3′

（2）PCR 反应体系 25μl 反应体积，含有 5～10ng DNA，1× PCR 缓冲液（10mmol/L Tris－HCl，pH 8.3，50mmol/L KCl，0.001％ 白明胶），1.5mmol/L $MgCl_2$，200μmol /L dNTPs，0.2μmol /L 每种引物，1.0u Taq DNA 聚合酶。

（3）PCR 热循环参数 93℃ 1min，58℃ 6min，循环 30 个周期，最后 70℃延伸 6min，退出循环。

11.3.3 D1S118 基因座

D1S118 基因座又称 p33.4、定位于 1 号染色体，是一个 GC 含量较高的基因座，Tm 值为 68℃，重复单位碱基数为 64bp，5′及 3′端侧翼碱基数分别为 11bp 和 96bp。

经对 100 个中国汉人调查，发现有 8 个等位基因（等位基因 7、10～15），在等位基因 13～14 之间发现一个片段长度介于两者之间的的罕见等位基因（表 11－6 的等位基因 14*）。等位基因片段长度为 603～1115bp，基因频率 0.005％～0.535％，等位基因频率如表 11－6，杂合度为 0.64，DP 值为 0.845。

表 11－6 P33.4 基因座等位基因频率（季安全）

等位基因	频率
7	0.03
10	0.01
11	0.03
12	0.19
13	0.535
14*	0.005
14	0.18
15	0.02
杂合度	0.64
DP	0.845

1. 引物序列 5′－CCG CGC CAG ACC CCA GCT GCT GAG－3′

5′－GCA GCA TAG GGC CTG TGC TGG GCT－3′

2. PCR 反应体系 50μl 反应体系，含有 50mmol/L KCl、10mmol/L Tris－HCl（pH 8.3），1.5mmol/L $MgCl_2$，0.001mg/ml 白明胶，200μmol/L dNTPs，0.2μmol/L 每种引物，1ng～1μg 基因组 DNA，95℃热变性 5min 后加入 1.25u Taq DNA 聚合酶，覆盖液体石蜡后置于 DNA 热循环仪中。

3. PCR 热循环参数 95℃变性 1min，72℃退火及延伸 4min，35 个循环结束后再在 72℃

延伸 7min。

11.3.4 D17S30 **基因座**

D17S30 又称 pYNZ22，是 Wolff 于 1988 年报道的一个 VNTR 基因座。它定位于人 17 号染色体短臂，由一个 70bp 重复单位构成，其核心序列为 TGG AGT CTC TGG GTG TGC GTC AGA CT，呈现高度多态性和个体特征。

在中国汉族中共发现 15 个等位基因 A1 ~ A15，以 123bp DNA 分子标准物或 ΦX174/HaeⅢ作为分子量标准，等位基因片段长度为 168 ~ 1148bp，杂合度为 0.82 左右，DP 值大于 0.94，等位基因命名没有采用国际通用命名法，而是根据片段大小按从小到大排列分型，将最小的等位基因命名为 A1，随后依次类推。各文章报道的群体数据表观上有所不同，详见表 11 - 7，这可能与调查的群体样本数目大小有关。在云南白族人群等位基因频率分布与汉族有显著差异（欧炯文等 1996）。

D17S30 基因座易发生较小等位基因优先扩增，尤其是当模板 DNA 量低时，易造成等位基因丢失。

表 11 - 7 pYNZ22 基因座各等位基因频率分布

等位基因	中国汉族（北京）	中国汉族	中国白族（云南）	西班牙人	意大利人
A1	0.100		0.1571	0.0365	0.088
A2	0.179		0.1000	0.1934	0.206
A3	0.304	0.0308	0.0524	0.1934	0.132
A4	0.142	0.0885	0.2190	0.2518	0.240
A5	0.108	0.078	0.1048	0.0438	0.064
A6	0.075	0.0192	0.1381	0.0511	0.054
A7	0.021	0.3538	0.0905	0.0146	0.005
A8	0.042	0.0346	0.0667	0.0255	0.024
A9	0.021	0.0885	0.0095	0.1058	0.064
A10	0.004	0.0769	0.0476	0.0620	0.083
A11	0.004	0.0615	0.0095	0.0182	0.024
A12		0.0731	0.0048	0.0036	
A13		0.0308			0.015
A14		0.0192			
A15		0.0514			
杂合度	0.73	0.8263	0.8778	0.8427	
DP 值	0.938	0.9494	0.9647	0.9453	

1. 引物序列 5′ - AAA CTG CAG AGA GAA AGG TCG AAG AGT GAA GTG - 3′

5′ - AAA GGA TCC CCC ACA TCC GCT CCC CAA GTT - 3′

2. PCR 反应体系 20μl 反应体积，含有每种引物 0.2μmol/L，模板 DNA 5 ~ 50ng，Taq DNA 聚合酶 1.2u，dNTPs 200μmol/L，$MgCl_2$ 1.6mmol/L，1 × PCR 标准缓冲液。

3. PCR 热循环参数 94℃ 30s，55℃30s，72 ℃ 60s，共 30 循环，最后 72 ℃ 7min，退

出循环。

或97℃ 1min，63℃ 1min，72℃ 6min，3个循环后 95℃ 1min，63℃ 1min，72℃6min，30个循环，最后72 ℃ 7min退出循环。

11.3.5 D1S111 基因座

D1S111 基因座又称 p 33.6，定位于人染色体 1 q21 - 1q31，是 Jeffreys 等人 1988 年首先报道的一个 VNTR 基因座，因用 33.6 探针检测发现该小卫星，故称为 p33.6 基因座（Loupart，1991）。核心序列为 37bp，5′端非重复碱基为 14bp，3′端非重复碱基为 45bp。已检出 14 个等位基因，核心序列重复数为 9～22，片段长度为 430～925bp，杂合度 0.76～0.72，DP 值 0.916～0.9456。在中国群体中检出 12 个等位基因，按照等位基因片段长度的顺序以 A1 - A12 命名，等位基因频率见表 11 - 8。

表 11 - 8 p33.6 基因座基因频率分布

等位基因	频率
A1	0.01
A2	0.165
A3	0.015
A4	0.02
A5	0.27
A6	0.02
A7	0.05
A8	0.02
A9	0.02
A10	0.03
A11	0.01
A12	0.02

1. 引物序列 5′ - TGT GAG TAG AGG AGA CCT CAC ATT - 3′

5′ - AGG TGA GAC ATT ACT CAA TCC AGG - 3′

2. PCR 反应体系 50μl 反应体系，含有 50mmol/L KCl，10mmol/L Tris—HCl（pH 8.3），1.5mmol/L $MgCl_2$，0.001mg/ml 白明胶，200μmol/L dNTPs，0.2μmol/L 每种引物，1ng～1μg DNA，95℃变性 5min 后，加人 1.25u Taq DNA 聚合酶，液体石蜡油覆盖后置于 DNA 热循环仪。

3. PCR 热循环参数 95℃变性 1min，60℃退火 1min，70℃延伸 4min，35 个循环后再 70℃延伸 7min。

11.3.6 COL2A1 基因座

COL2A1 基因座是人类胶原蛋白Ⅱ型基因（collagen Ⅱ gene）（GDB：9862018），定位于 12q14.3，是该基因 3′端富含 AT 区域的一个 VNTR 基因座，重复单位 31～34bp，已检出 10 个等位基因，陈松等（1996）在中国汉族人群检出 7 个等位基因，等位基因片段长度为 600～900bp，等位基因频率为 0.025～0.395，杂合度 0.697，DP 值 0.845。

1. 引物序列　5′-CCA GGA TAA GGT TCA CAT CT-3′

5′-GTC ATG AAC TAG CTC TGG TG-3′

2. PCR 反应体系　50μl 反应体系，含有 50mmol/L KCl、10mmol/L Tris-HCl（pH8.3），1.5mmol/L $MgCl_2$，0.001mg/ml 白明胶，200μmol/L 的 dNTPs，0.1μmol/L 的每种引物，1ng ~ 1μg DNA，95℃热变性后，加入 1.25u DNATaq 聚合酶。

3. PCR 扩增循环参数　95℃变性 1min，62℃（或 55℃）退火 1min，72℃延伸 4min，35 个循环后 72℃延伸 7min。

11.3.7 IL-6 基因座

IL-6 基因座是人类白细胞介素 6 基因（Interleukin - 6 gene），3′端富含 AT 区域的一个 VNTR 基因座，定位于 7p15-p21。在中国汉族人群中已检出 5 个等位基因，等位基因片段长度为 600~800bp，等位基因频率为 0.03~0.5905，杂合度 0.592，DP 值 0.791。

1. 引物序列　5′-GCA ACT TTG AGT GTG TCA CG-3′

5′-GCA GTG ATG GAT GCA ACA C-3′

2. PCR　反应　50μl 反应体系，含有 50mmol/L KCl、10mmol/L Tris-HCl（pH 8.3），1.5mmol/L $MgCl_2$，0.001mg/ml 白明胶，200μmol/L dNTPs，0.1μmol/L 引物，1ng ~ 1μg DNA，95℃热变性后加入 1.25u TaqDNA 聚合酶。

3. PCR 扩增循环参数　95℃变性 1min，55℃退火 1min，72℃延伸 4min，35 个循环后 72℃延伸 7min。

11.3.8 D4S95 基因座

D4S95 基因座定位于人类第 4 号染色体，与 Huntington 疾病有关。在中国汉族人群中已检出 7 个等位基因，18 种不同的基因型，片段长度为 910 ~ 1156bp，重复序列长度为 39bp，基因频率为 0.005 ~ 0.356。

1. 引物序列　5′-GCA TAA AAT GGG GAT AAC AGT AC-3′

5′-GAC ATT GCT TTA TAG CTG TGC CTC AGT TT-3′

2. PCR 反应体系　50μl 反应体系，含有 50mmol/L KCl、10mmol/L Tris-HCl（pH 8.3），1.25mmol/L $MgCl_2$，0.025% BSA，200μmol/L dNTPs，1.25μmol/L 两种、1ng ~ 1μg DNA，95℃热变性后加入 1.25u DNA Taq 聚合酶。

3. PCR 扩增循环参数　95℃变性 1min，60℃退火 2min，72℃延伸 3min，30 个循环。

11.3.9 DXS52 基因座

DXS52 基因座定位于人类染色体 Xq28，与 A 型血友病基因（hemophilia A）紧密连锁，由 Oberle 等人于 1985 年发现。重复单位长约 60bp，侧区序列长约 650bp。已检出 16 个等位基因（表 11-9）。Y 染色体上无 DXS52 基因座，男性只有一个 DXS52 基因，DNA 的扩增产物只有一条带。女性有两个 DXS52 基因，扩增产物可出现一条带（纯合子），也可出现两条带（杂合子）。中国汉族女性的杂合度为 0.77，DP 值 0.93，在男性中 DP 值为 0.89。

表 11－9 DXS52 基因座等位基因频率

等位基因	频率	
	中国汉族	美国白人
1	0.171	0.120
4	0.051	0.000
5	0.025	0.000
8	0.006	0.000
9	0.063	0.002
10	0.120	0.020
11	0.221	0.100
12	0.089	0.000
13	0.051	0.000
14	0.070	0.140
15	0.025	0.020
16	0.051	0.360
18	0.025	0.000
29	0.032	0.120
37	0.000	0.080
39	0.000	0.020

1. 引物序列 5′－GCC ATG TCA TCACTT CTC TCA TGT T－3′

5′－CAC CAC TGC CCT CAC GTC ACT T－3′

2. PCR 反应体系 20μl 反应体系，含有 50mmol/L KCl、10mmol/L Tris－HCl（pH 8.3），1.5mmol/L $MgCl_2$，0.001mg/ml 白明胶，200μmol/L 的 dNTPs，20pmol 的引物，10～20ng DNA，0.5u Taq DNA 聚合酶，5%甲酰胺。

3. PCR 扩增循环参数 94℃变性 1min，58℃退火 2min，72℃延伸 3min，30 个循环。

11.4 小卫星 VNTR 扩增分析应用中注意事项

扩增片段长度多态性分析应用 PCR 扩增技术将某一 VNTR 基因座的等位基因片段扩增出来，然后通过电泳技术将这些扩增出来的片段按大小分离开来。为保证结果的准确，在检验中必须注意 PCR 反应中涉及的每一环节和因素。

11.4.1 污染问题

污染问题是 PCR 技术的一个共性问题，包括样品 DNA 污染以及反应体系试剂污染等，将严重干扰 PCR 分析，详见 PCR 章节。

11.4.2 模板 DNA

模板 DNA 的量和质影响 PCR 扩增。模板量太少，得不到扩增产物；模板量太多，扩增特异性会降低，非特异带影响结果的判读。小卫星 VNTR 扩增片段相对较大，灵敏度较 STR 低，一般模板 DNA 量 1～100ng 为宜，在进行 PCR 反应前最好进行 DNA 定量。

有时也会出现因模板 DNA 太多反而得不到扩增，主要原因是模板 DNA 不纯，含有一些 Taq DNA 聚合酶抑制剂，如斑痕检材 DNA 含有载体的染料、血痕中的血红素等。稀释

样品 DNA 浓度可降低抑制物的浓度，取得良好效果。如果在一系列浓度比例下仍得不到扩增，说明模板 DNA 液中含太多的聚合酶抑制物，需要纯化模板 DNA。

11.4.3 较小等位基因优先扩增

由于小卫星 VNTR 的串联重复单位相对较大，等位基因片段间长度相差很大，易产生较小等位基因优先扩增现象，二个等位基因扩增产量差异明显，易将杂合子误判为纯合子，尤其是 D17S30，ApoB 基因座，这种现象更为常见。模板 DNA 用量越少，越容易出现等位基因丢失。当微量检材的检验结果为一条谱带的纯合子时，分型时要加以注意，以免造成错判。PCR 反应前进行模板 DNA 定量，以避免扩增中较大等位基因丢失的可能性。

11.4.4 降解 DNA 样品大等位基因的丢失

轻度降解的模板 DNA 一般不会影响 VNTR 基因座的检验，但是，如果 DNA 严重降解，较大的等位基因就可能不被扩增，使杂合子样品呈现为纯合子型。建议在进行 PCR 反应前用琼脂糖凝胶电泳检查模板 DNA 分子的质量，观察是否存在降解以及降解的程度，以利对结果的解释。

11.4.5 试剂和仪器质量及性能

PCR 引物、Taq DNA 聚合酶等试剂及扩增仪的质量和性能必须符合分型检测的技术标准。

不同厂家的 Taq DNA 聚合酶的活力可能不同，换用不同厂家或不同批号的酶要做预实验。Taq DNA 聚合酶要求在 -20℃ 低温保存，贮存浓度为 5u/μl，稀释后的低浓度酶保存一段时间后易失活。引物反复冻融，易引起降解，引物应分装冷冻保存。进行 PCR 反应时，要求设置已知分型结果的样品作阳性对照，以检测反应系统是否准确有效。

不同型号的热循环仪，控温机制不同，温度变化速度不同、应定期测试与校正，以达到最佳扩增效果。

第十二章 微卫星（STR）基因座概述及检验方法

12.1 STR 基因座概述

12.1.1 STR 基因座特征

重复单位长度为 2～6bp 的重复序列称为微卫星序列，也叫短串联重复序列(short tandem repeat,STR),或简单序列重复(simple sequence repeats,SSRs),统称 STR。STR 基因座重复序列串联排列，基因座扩增片段一般在 400bp 以下。STR 基因座更适合于 PCR 扩增而没有优先扩增的问题，因为在 STR 基因座杂合子个体的二个等位基因大小比较相近。目前，STR 是法医 DNA 分析中最常用的一类遗传标记。STR 在整个人基因组分布广泛，平均每 10kb 出现一个。多态 STR 序列绝大多数位于非编码区，极少数在编码区域。目前在人基因组 DNA (23 对染色体)发现的 STR 序列达 8000 多,部分法医 DNA 分析应用的基因座见表 12－1。

表 12－1 部分 STR 基因座

染色体	STR 标记
1	F13B, RENA4, D1S1171, D1S1650
2	TPOX, D2S410, D2S436, D2S1242, D2S441
3	D3S1349, D3S1352, D3S1358, D3S1359 , D3S1744, ACPP
4	FGA, FABP, GABARB15,
5	CSF1PO, D5S373, D5S818,
6	F13A01, ACTBP2, FOLP23, D6S366, D6S502, D6S965,
7	D7S460, D7S809, D7S820, D7S1517, D7S1520, D7S2846, D7S2201
8	LIPOL, D8S306, D8S320, D8S323, D8S344, D8S347, D8S369, D8S1179, D8S1132
9	D9S52, D9S304
10	D10S89, D10S2325
11	TH01, APOAⅡ, D11S554, UGB
12	VWA, CD4, PLA2A1, D12S67, D12S391, D12S1090
13	D13S308, D13S317
14	D14S306
15	FES/FPS, CYAR04
16	D16S539, D16S537
17	D17S976
18	MBP, D18S51, D18S535, D18S849
19	D19S253, D19S400
20	D20S85, D20S470
21	D21S11
22	D22S686, D22S533, D22S685, D22S683, D22S445, D22S444
X	HPRTB, ARA, STRX1, DXS6807
Y	DYS19, DYS385, DYS389 Ⅰ, DYS389 Ⅱ, DYS390, DYS391, DYS392, DYS393, DYS156Y, A10, C4, 7.1, 7.2, A4, H4, DYS434, DYS435, DYS436, DYS437, DYS438, DYS439

STR 基因座具有以下特点：

1.STR 基因座在人类基因组中分布广泛，多态性基因座数量多。

2.STR 基因座等位基因片段长度一般小于 400bp，易于扩增。对 DNA 的质量要求低，适于陈旧、降解生物检材的 DNA 分析。

3.STR 分型检测的灵敏度比小卫星 VNTR 基因座高 10 倍，所需的检材 DNA 量更少，适用于微量、超微量检材的鉴定。

4.STR 基因座等位基因片段长度差异较小，采用高分辨率 PAGE 可以获得不连续基因频率分布参数。

5.STR 基因座等位基因之间长度差异较小，较小等位基因优先扩增现象不明显。虽然某些基因座可以观察到杂合子中二个等位基因信号强度差异，但二者强弱差异小于 40%。

6.STR 基因座的等位基因片段比较小，各个基因座扩增条件相似，因此多个 STR 基因座可以安排在同一个反应体系中进行复合扩增，简化操作程序，节省检材与试剂，提高工作效率。

7.STR 基因座分析方法简便、判型准确，分型程序明确、规范，所得分型结果图形简单，有利于实现 DNA 分型标准化和自动化，有利于 DNA 分型数据的计算机存储、联网检索。

12.1.2 STR 基因座的分类与等位基因命名

1.STR 基因座的分类

确定新的 STR 基因座有两个方法：（1）在 DNA 序列数据库如 GenBank 中搜寻有 6 个以上连续重复单位的基因座；（2）用分子生物学分离方法。

STR 基因座常根据重复单位的核苷酸数分类。二核苷酸序列（dinucleotide repeat）指有不间断重复的二个核苷酸重复，三核苷酸（trinucleotides）指的是重复单位为三个核苷酸，四核苷酸（tetranucleotide）、五核苷酸（pentanucleotide）和六核苷酸（hexanucleotide）分别指 4、5 和 6 个核苷酸为重复单位序列。理论上，单、二、三、四、五和六核苷酸重复有 4、16、64、256、1024 和 4096 种重复单位类型，又称主型（motif）。但由于微卫星是串联重复，实际的主型数要比理论的少许多，因为有些主型表面上看起来不同，实质上是相同的。比如，主型（GAAA）n 相当于（AGAA）n 或（AAGA）n、（AAAG）n 或（TTTC）n、（TTCT）n、或（TCTT）n 或（CTTT）n，这 8 个主型其实指的是同一个序列，只是表达的方式不同。因此实际上，单、二、三、四、五和六核苷酸重复只有 2、4、10、33、102 和 350 个主型（表 12－2）。

表 12－2　不同重复单位长度的 STR 序列主型

	主型
单核苷酸序列	A、C
二核苷酸序列	AC、AG、AT、CG
三核苷酸序列	AAC、AAG、AAT、ACC、ACG、ACT、AGC、AGG、ATC、
四核苷酸序列	AAAC、AAAG、AAAT、AACC、AACG、AACT、AAGC、AAGG、AAGT、AATC、AATG、AATT、ACAG、ACAT、ACCC、ACCG、ACCT、ACGC、ACGG、ACGT、ACTC、ACTG、AGAT、AGCC、AGCG、ATCR、AGGC、AGGG、ATCC、ATCG、ATGC、CCCG、CCGG

四核苷酸序列基因座是法医 DNA 分析最常用的基因座，其中 AGAT 或 GATA 为其最常见的主型。

STR 序列不仅在重复单位的长度和重复数目上有差异，而且重复单位排列方式不同。根据重复单位的碱基组成，Urquhart（1994）等将 STR 分成以下三类：（1）简单重复序列（simple repeats）：含有同一长度和序列的重复单位，如 HumFES/FPS 基因座只含有单一的（ATTT）n 重复。HumCD4，TH01，F13B 等基因座也是简单重复；（2）复合重复序列（compound repeats）：含有二个或二个以上简单重复序列，如 HumGABRB15 基因座为（GATA)$_{5-12}$（GATC)$_{2-4}$（TATC)$_{1-2}$，HumVWA 基因座为（TCTA）n（TCTG）m（TCCA）p；（3）复杂重复序列（complex repeats）：同一基因座重复单位既有序列差异，又有长度差异，如 Hum ACTB、D21S11 基因座。

2.STR 基因座等位基因命名

为了保证实验室之间的数据重复性和可比性，国际法医血液遗传学会（International Society of Forensic Haemogenetics，ISFH）在 1994 年和 1997 年对等位基因命名提出了命名原则，统一了等位基因命名。1994 年的建议主要对等位基因分型标准物（allelic ladder）和含不完整重复序列等位基因的命名。1997 年的建议主要讨论的是 STR 序列和重复单位的命名。1997 年以前的文献中重复单位命名与 1997 年 ISFH 提出的原则不同，如果不加以注意，可能会引起一些混淆。例如，早期英国法庭科学服务部（FFS）将 TH01 基因座的核心序列标为 TCAT，而 Caskey 等将 TH01 的核心序列表达为 AATG（Edwards，1991）。同一个 STR 基因座，出现有两种不同的核心序列，其主要原因是不同的作者采用不同的 DNA 链进行命名。DNA 分子是一个双链分子，都可以按 5′→3′读序，链的选择决定着序列的命名。例如某一基因座的一条链的序列为 5′...（TCAT）n...3′，则同一序列的另一条互补链的则为 5′...（ATGA）n...3′，选择不同的链命名，核心序列的命名就不同，而且根据重复区域周围的序列，另一条的核心重复序列有可能会前后发生一些移动，如图 12－1 所示，按上面一条链的序列，其核心序列为 TCAT，而按下面的则为 TGAA，向 5′端移动了 3 个碱基。

```
               1     2     3     4     5     6
5′-TTTCCC  TCAT  TCAT  TCAT  TACT  TACT  TACT  TCACCATGGA-3′
3′-AAAGGG  AGTA  AGTA  AGTA  AGTA  AGTA  AGTA  AGTGGTACCT-5′
                 6     5     4     3     2     1
```

图 12－1 STR 基因座的序列

因此必须采有统一的命名原则。ISFH 1997 年提出的命名原则如下：

（1）链的选择

①对于蛋白编码区域的 STR 基因座（包括基因的内含子）选用编码链作为重复单位命名链。如 VWA、TPOX、CSF1PO 基因座遵循此原则。

② 与蛋白质编码无关的如以 D＊S＊＊命名的 STR 基因座，采用第一个进入公共数据库的原始序列作为标准进行命名。

③ 命名已建立，但与原则不一致的命名仍按原先的命名，以免引起混淆，如 TH01 以“AATG”为核心序列。

(2) 主型的选定和等位基因命名

①重复序列主型从离 5′端最近的一个核苷酸开始定义。例如：

5′ - GG TCA TCA TCA TGC - 3′序列可以看作有 3 × TCA 的重复序列，也可以看作 3 × CAT 的重复序列。但根据 ISFH 新原则，只有第一个（3 × TCA）是正确的，因为是从离 5′端最近的一个核苷酸开始。

② 根据等位基因内所含的重复单位数目命名等位基因，如某一等位基因含有 20 个重复序列，则该等位基因命名为 20。

③ 含有不完整重复序列主型的等位基因命名：在完整等位基因数后面加一小数点，小数点后面为不完整重复序列含有的碱基数。例如，TH01 基因座中含有一个不完整重复主型的等位基因，等位基因 9.3，表示含有 9 个完整的四核苷酸重复 AATG 和一个不完整的 3 个核苷酸的 ATG 重复。又如 D21S11 基因座的 17.2 等位基因，表示含有 17 个四核苷酸重复单位和一个含 2 个核苷酸的不完整重复单位。

④ 经测序确定的并按 ISFH 建议命名的等位基因分型标准物（ladder）为分型标准，进行未知样品的等位基因分型。

3. 等位基因分型标准物制备

等位基因分型标准物（allelic ladder）是人群中常见的已知等位基因分型的等位基因混合物，用来与未知样品比对确定基因型。通过大量群体调查，筛选、分离出这些等位基因，经 DNA 测序确定这些等位基因的片段长度与序列（重复序列重复次数），明确等位基因分型。将这些等位基因混合起来再进行扩增，制成含有各种等位基因的分型标准物。

也可以将等位基因克隆到质粒里，进行制备。

12.1.3 微变异体和等位基因分型标准物以外的等位基因

1. 微变异体等位基因的定义

在大规模人群调查中，发现存在一些大小与等位基因标准物只差一个或几个碱基的稀有等位基因，这是由于碱基的插入、缺失等造成的。这些稀有等位基因与完整重复单位的等位基因差异极其微小，因此称为微变异体（microvariant）。由于微变异体的大小与作为对照的等位基因分型标准物不同，大部分不包含在分型标准物中，因此将微变异体也称为分型标准物以外的等位基因（off - ladder allele）。

最常见的微变异体的例子是 TH01 基因座的等位基因 9.3，它含有 9 个完整的重复单位（AATG）和一个 3 个碱基（ATG）的部分重复单位。等位基因 9.3 与等位基因 10 的差异是在第 7 个重复单位发生一个腺苷酸残基的缺失。

随着调查样本数的不断增加，越来越多的微变异体和稀有等位基因被发现，Crouse（1999）调查 10 000 个样本发现在 CSF1PO，TPOX 和 TH01 基因座有 42 个微变异体存在，张纯斌（2000）分析 1 772 个样本，在 9 个 STR 基因座中（AmpFISTR Profiler Plus）发现 51 个微变异体（表 12 - 3），季安全（2000）报道在 FGA 基因座检出等位基因 13，14，15 和 16，在 D18S51 基因座检出等位基因 7，15.2，17.1，27 和 28，在 D13S317 基因座检出等位基因 5，6 和 7，D3S1358 基因座检出 10，11，20，在 D21S11 基因座等位基因 22.3 相对常

见，在 D7S820 基因座则为等位基因 9.1 较常见。Shewale（2000）报道在 7 000 个样品中有 24 个样品出现稀有等位基因，各等位基因命名和出现频率如表 12－4 所示。微变异体多发生在那些含有复杂重复单位结构的高度多态性的基因座，比如 FGA，D21S11 和 D18S51 基因座。

表 12－3　1772 例无关中国汉族个体中稀有等位基因出现情况

基因座	等位基因	观察到的个体数	出现频率（%）
D21S11	30.1	11	0.620
FGA	17	6	0.339
	16	4	0.225
	14	1	0.056
	13	1	0.056
D7S820	9.1	17	0.959
	9.2	1	0.056
	10.1	6	0.339
	13.3	1	0.056
D18S51	27	2	0.113
D13S317	7	1	0.056

表 12－4　7000 例个体中 Profiler Plus 系统 STR 基因座的稀有基因频率

基因座	等位基因	观察到的个体数	出现频率（%）
D3S1358	14.3	1	0.0143
FGA	42	2	0.0285
	43	2	0.0285
	44.1	1	0.0143
	45.1	2	0.0285
	16.1	1	0.0143
	47.1	1	0.0143
	48.1	1	0.0143
D21S11	33.1	2	0.0285
D18S51	11.2	1	0.0143
	21.3	1	0.0143
	22.2	1	0.0143
D5S818	12.3	1	0.0143
D7S820	7.3	1	0.0143
	9.1	2	0.0285
	10.3	1	0.0143
	11.1	2	0.0285
	12.1	1	0.0143

2. 微变异体等位基因的确定与命名

如何确认微变异体？一般在杂合子个体中发现的微变异体等位基因比较容易确认，一个等位基因的大小与 ladder 的等位基因一致，而另一个相差较大，如图 12-2 所示，样品在 FGA 基因座含有一个等位基因 25 和标有“off-ladder allele”的第二个峰，它位于等位基因 28 和 28.2 之间。当一个样品出现“off-ladder allele”时，首先核实该样品的分子量内标是否准确，检查样品等位基因与 ladder 中相应等位基因的片段大小，计算各个等位基因样品与 ladder 的大小差异 $\delta 1$ 与 $\delta 2$（$\delta 1 = -0.12$bp，$\delta 2 = +0.87$bp）。根据 $\delta 1$ 与 $\delta 2$ 计算杂合子个体的二个等位基因之间的相对漂移 0.99 bp（$\delta 1 - \delta 2$），此“off-ladder allele”比等位基因 28 大 1 bp，那么它就是 FGA 基因座的一个真正的等位基因 28.1。

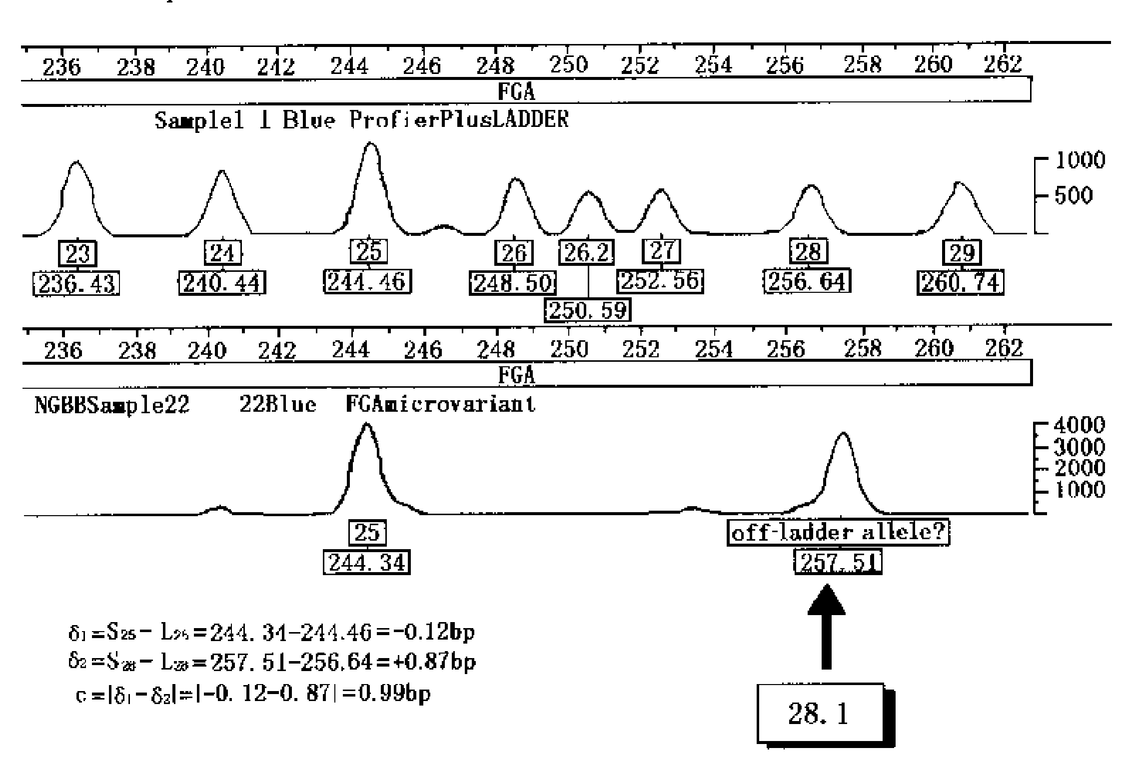


图 12-2　微变异体等位基因的检测

当稀有等位基因出现在纯合子样品时，不能按上面介绍的杂合子个体中出现的情况处理，该稀有等位基因只能与 ladder 中相近峰的片段长度比较，进行命名。

像上面介绍的微变异体等位基因出现在等位基因分型标准物之间，以相邻等位基因名后加“.x”命名此类微变异体等位基因。如图 12-2 中，命名为“28.1”。然而，有时一些稀有等位基因出现在等位基因分型标准物两端之外，或小于最小的等位基因，或大于最大的等位基因，此时无法确定这些稀有等位基因的确切命名，只能命名为“<”或“>”ladder 中靠得最近的等位基因。

对于复合扩增系统，有时微变异体发生在二个基因座之间，此时要求用单个基因座特异的引物进行单个扩增，确认微变异体等位基因。

12.1.4 长度相同但序列不同的等位基因

一些复杂重复单位序列的基因座，例如 D21S11，VWA 基因座，含有不同重复单位，等位基因的长度相同但重复单位内在的排列顺序发生变化。如 D21S11 基因座有 4 个长度为 210bp 的等位基因，它们的序列结构分别为

$[TCTA]_4$ - $[TCTG]_6$ - CR - $[TCTA]_{12}$

$[TCTA]_5$ - $[TCTG]_6$ - CR - $[TCTA]_{11}$

$[TCTA]_6$ - $[TCTG]_5$ - CR - $[TCTA]_{11}$

$[TCTA]_6$ - $[TCTG]_6$ - CR - $[TCTA]_{10}$

其中，CR 表示为恒定区序列。这种变异的等位基因只有通过对等位基因直接测序才能检测出，一般情况下不能区分开它们。

12.1.5 三条带模式

在单一个体的某个基因座检测出 3 条带或峰，并具有重现性，显示一个基因座有 3 个“等位基因”。造成这种情况的原因目前尚不清楚，推测是存在额外的染色体或引物结合点。

三条带模式的 3 个峰的信号强度可以相近，也可以差异悬殊。比如 Crouse 等（1999）人发现的 TPOX 基因座的 3 个峰信号强度比较一致，而作者在 VWA 基因座和 D18S51 基因座发现的 3 个峰的信号强度不等，其中 2 个峰的信号强度比较接近（图 12-3）。目前已有报道在 TPOX，CSF1PO，FGA，D7S820，D5S818，D21S11，D16S539，VWA 和 D18S51 等基因座出现 3 条带模式。CODIS 系统 13 个 STR 基因座出现三条带的百分率见表 12-5。

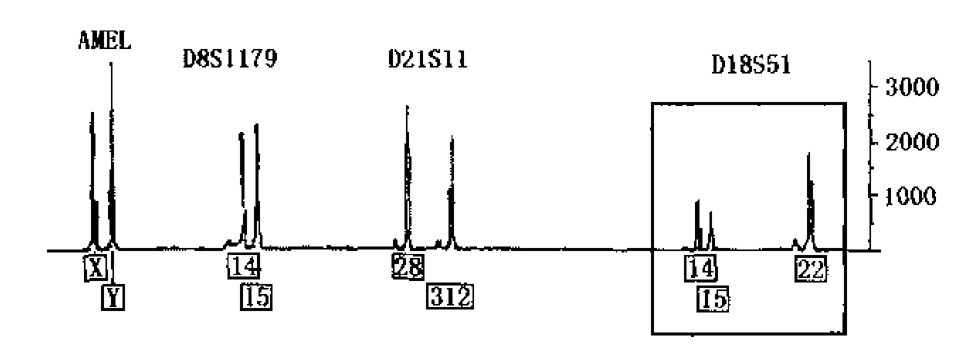


图 12-3 D18S51 基因座的三条带图谱

表 12-5 CODIS 系统 13 个 STR 基因座三条带和无效等位基因出现百分率（美国血库协会 1999）

基因座	多带出现率（%）	无效等位基因（%）
CSF1PO	未报道	2/42020（<0.01）
FGA	未报道	2/1104（0.18）
TH01	0/2646（<0.040）	2/7983（0.03）
TPOX	13/42020（0.03）	11/43704（0.03）
VWA	1/6581（0.02）	7/4220（0.02）
D3S1358	未报道	未报道
D5S818	未报道	3/74922（<0.01）

续表

D7S820	1/406（0.25）	1/42020（<0.01）
D8S1179	未报道	未报道
D13S317	未报道	52/62344（0.08）
D16S539	0/1165	3/52959（<0.01）
D18S51	未报道	未报道
D21S11	未报道	1/203（0.49）

12.1.6 等位基因的丢失和无效等位基因

在STR基因座的重复单位内部、旁侧区和引物结合区（图12-4）序列中出现一些碱基序列的变异，可能会影响PCR扩增。最严重的情况是，发生等位基因丢失现象（allele dropout），只有一个扩增片段，原本杂合子个体变成了一个“纯合子”。

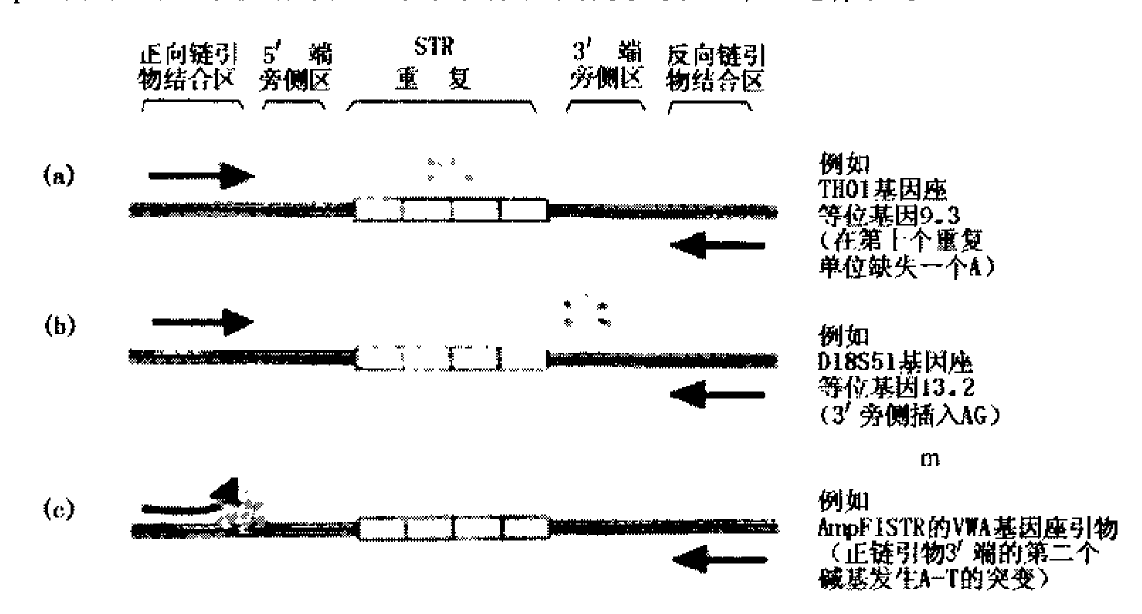


图12-4　STR基因座重复区内或侧翼序列变异的三种情况

*代表此处DNA序列不同（插入、缺失或碱基变换）

(a) 变异发生在重复区内，对引物结合没有影响，PCR扩增几乎不受影响，只是等位基因大小有差异；(b) 变异发生在重复区外的旁侧，但在引物结合点的内侧，尽管扩增产物大小有变化，但对PCR扩增没有影响；(c) 变异发生在引物结合区，由于引物与模板不再完全互补，引物不能退火，不能进行PCR扩增。

此时碱基变异正好发生在模板DNA的PCR引物结合区，引物不能退火，不能扩增，没有扩增产物。模板DNA中虽然存在这个等位基因，只是由于引物不能退火使扩增失败，称之为无效等位基因（null allele）。STR重复单位侧翼序列相对较稳定，出现这种无效等位基因的现象较少。表12-5列出了美国血库协会1999年公布的CODIS系统13个STR基因座出现无效等位基因的百分率。

1. 无效等位基因的发现

比较不同序列引物的扩增结果可以发现无效等位基因。例如，Kline（1999）用

Promega 公司的 Powerplex 试剂盒和 PE Applied biosystem 公司的 AmpFLSTR Blue 试剂盒分析 600 例个体，分析比较两个试剂盒的 VWA 基因座结果，发现有一个样品用 Powerplex 试剂的分型结果为 16，19，而 AmpFLSTR 的则为 16，16。VWA 基因座的等位基因 19 丢失了，因为该个体在正向链引物的 3'端有序列变异，在引物 3'端或 3'端附近碱基发生突变，使 PCR 扩增过程中很少延伸或根本没有延伸，最终导致等位基因的丢失。像上面提到的例子中样品拥有等位基因 19，实质上是一个杂合子个体，只是使用 AmpFLSTR 的引物时没有得到扩增。后来测序结果证实是由于模板 DNA 在 AmpFLSTR 的 VWA 基因座的正向链引物的 3'端的第二个核苷酸发生了 A - T 的变化导致无效等位基因的发生。AmpFLSTR 的 VWA 引物序列与 Powerplex 的不同。检测结果明显不同。

通过对实际观察到的纯合子数与由 Hardy - Weinberg 平衡计算的纯合子的期望值比较，可以预测是否存在无效等位基因。如果纯合子数不正常的高，暗示可能存在无效等位基因。因此，在新的 STR 基因座用于检验前必须认真进行群体调查，获得人群数据。

2. 无效等位基因的解决办法

如果一个基因座检测到一个无效等位基因，可能的解决办法如下：

(1) 重新设计引物，避开有变异的碱基位置。但可能导致新引物序列干扰复合体系，需要重新优化 PCR 反应条件。这比较费时，工作量大，不是一个十分理想的方法。

(2) 放弃这个基因座。这个方法在研究建立复合体系时是较理想的方法，英国 FSS 发现 D19S253 存在无效基因后，他们在开发的第二代复合体系中就去除了 D19S253 基因座。

(3) 反应体系中加一个已知序列多态性的“degenerate”引物，这个增加的引物能够扩增引物结合位置突变的等位基因。如果引物结合点的序列变异非常稀有，那么不值得在复合体系中加一个额外的等位基因。

(4) 对于引物结合位置发生有变异的样品，降低退火温度，重新扩增该样品。如果引物只是稍微不稳定，较低退火温度可获扩增产物，可以观察到杂合子样品的峰高不平衡（图 12 - 5），通过降低 PCR 过程的退火温度可以纠正峰高的不平衡。

没有一对引物可以绝对地说不会产生无效等位基因现象。在个体识别鉴定，只要用同一套引物、在相同条件分析现场物证与比对样品，如果相同，说明它们来自于同一个体，无须考虑无效等位基因问题。如果建立多实验室的数据库，进行联网比对，那么潜在的等位基因丢失可能会导致假排除。采用统一的试剂或者降低数据库搜索的标准，可以克服这个潜在的问题，例如允许有一个基因座的不完全匹配。但对于亲子鉴定，会出现假排除现象。

在选择引物进行 STR 基因座扩增时，需仔细评估候选引物，尽可能地避开引物结合点突变的引物序列。具体做法有：①测序分析多个引物，观察是否有序列多态性；②进行家系遗传调查，亲代与子代基因传递是否符合孟德尔定律；③检查杂合子样本的二等位基因峰信号是否平衡，降低退火温度重新扩增表观的纯合子样本，是否有纯合子变成了杂合子；④群体调查，分析纯合子数观察值与期望值间的偏差。避免无效基因的关键是在设计复合扩增 STR 引物时，保证引物结合点落在保守序列区内，而又不干扰其他基因座的扩增。

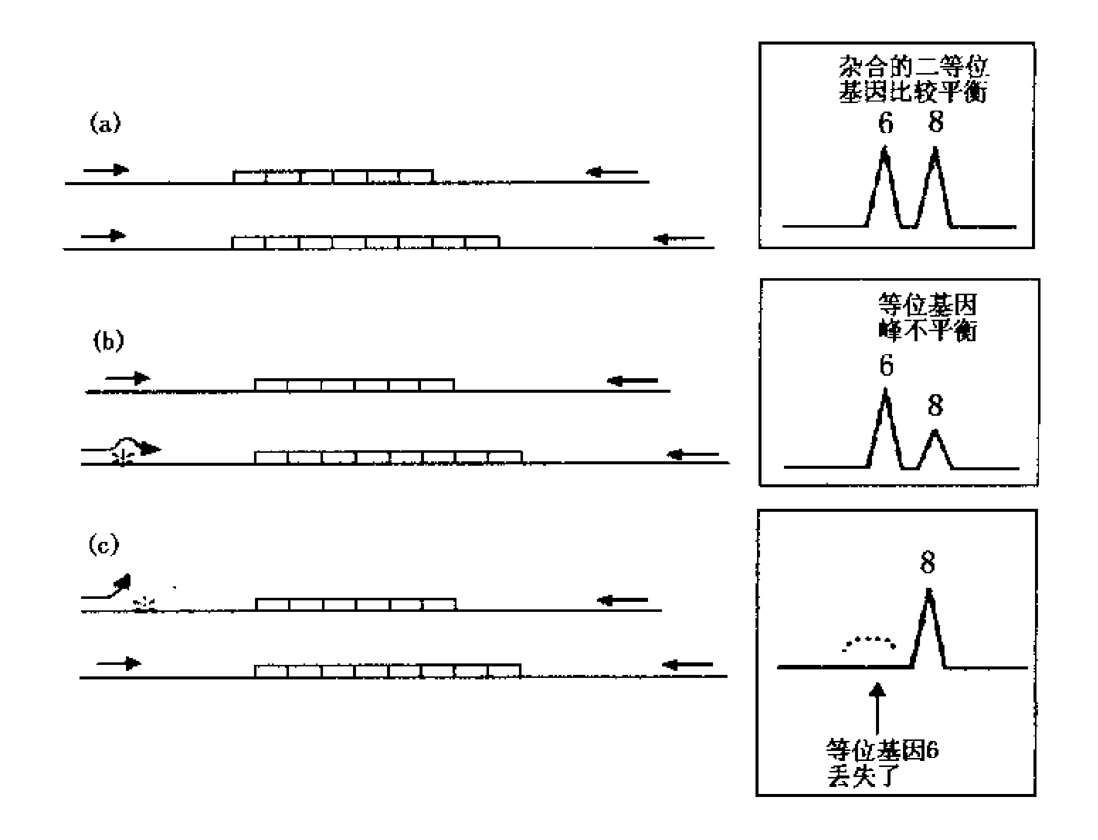


图 12-5　杂合子个体的引物结合处的序列多态性对 PCR 扩增的影响

(a) 杂合子的二个峰平衡 (b) 峰不平衡 (c) 一个等位基因丢失

12.1.7 突变

STR 基因座的任何 DNA 序列都可能发生突变。理论上，迄今存在的所有等位基因都是最早极少数个体的等位基因经过成千上万年的进化过程缓慢变化的结果。突变可以是单个碱基的变化，也可以是整个重复单位长度变化。STR 突变的机制主要是复制滑动（replication slippage）或 DNA 复制的错误修复。据文献估计 STR 基因座平均突变率为 0.1% ~ 0.5%。

1. 等位基因突变的发现

根据孟德尔遗传定律，子代的遗传物质一半来自于母亲、一半来自于父亲，检查母、子、父三联体的基因型，如果孩子的等位基因与父母的有差异，表明有可能存在突变（图 12-6）。由于大多数 STR 基因座的突变率比较低，必须调查大量的亲代-子代间等位基因的传递，才能发现突变。国际 DNA 委员会建议在新 STR 基因座调查中至少要观察 500 次减数分裂。一般来自父方的突变率高于母方 5 倍以上（17:3）。

2. 突变基因的确定

Brinkmann（1998）根据 Weber（1993）报道的文献，归纳如何确定被突变的等位基因。如果存在有两种可能的等位基因发生突变，那么距突变成的新等位基因长度相差最小的等位基因就是发生突变的基因。例如，一个父方等位基因与子代等位基因差一个重复单位，而另一个差 2 个重复单位，则可以推断与子代等位基因只相差一个重复的那个等位基因发

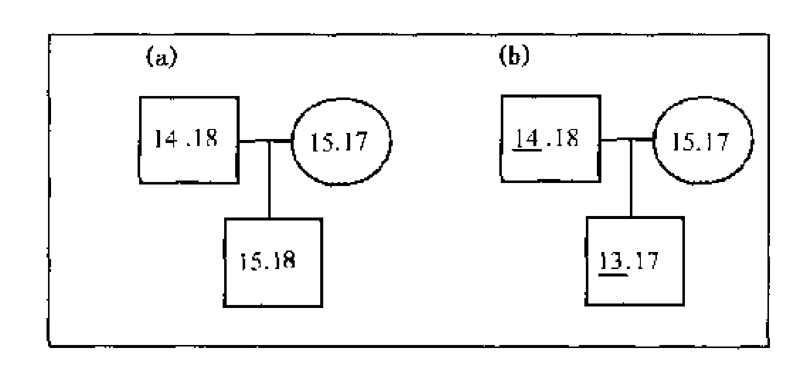


图 12-6 三联体家系中观察的的突变

(a) 正常等位基因传递模式 (b) 父方等位基因 14 发生突变变成为等位基因 13

生了一步突变（one step mutation）。如果与子代等位基因比较，二个亲代等位基因的差别一样，就无法确定哪个等位基因发生了突变。

大多数 STR 基因座的突变涉及单个重复单位的增加或减少，即一步突变，大约占突变基因的 90%。例如图 12-6 中 VWA 基因座，突变以一步突变为主，即单个重复单位的增加或减少为主，父方有 14 个重复单位的等位基因，由于突变，子代表现为 13 或 15 个重复单位。突变使重复单位增加与减少的比例为 2:1，因此总体上突变使链延长（Wiegand，2000）。Brinkmann（1998）观察了 9 个 STR 基因座 10 844 次突变发现了 23 个突变，在 23 个突变中有 22 个为一步突变，另一个为二步突变，没有观察到不完整重复单位的插入或缺失。图 12-7 为发现的 23 个突变情况示意图。

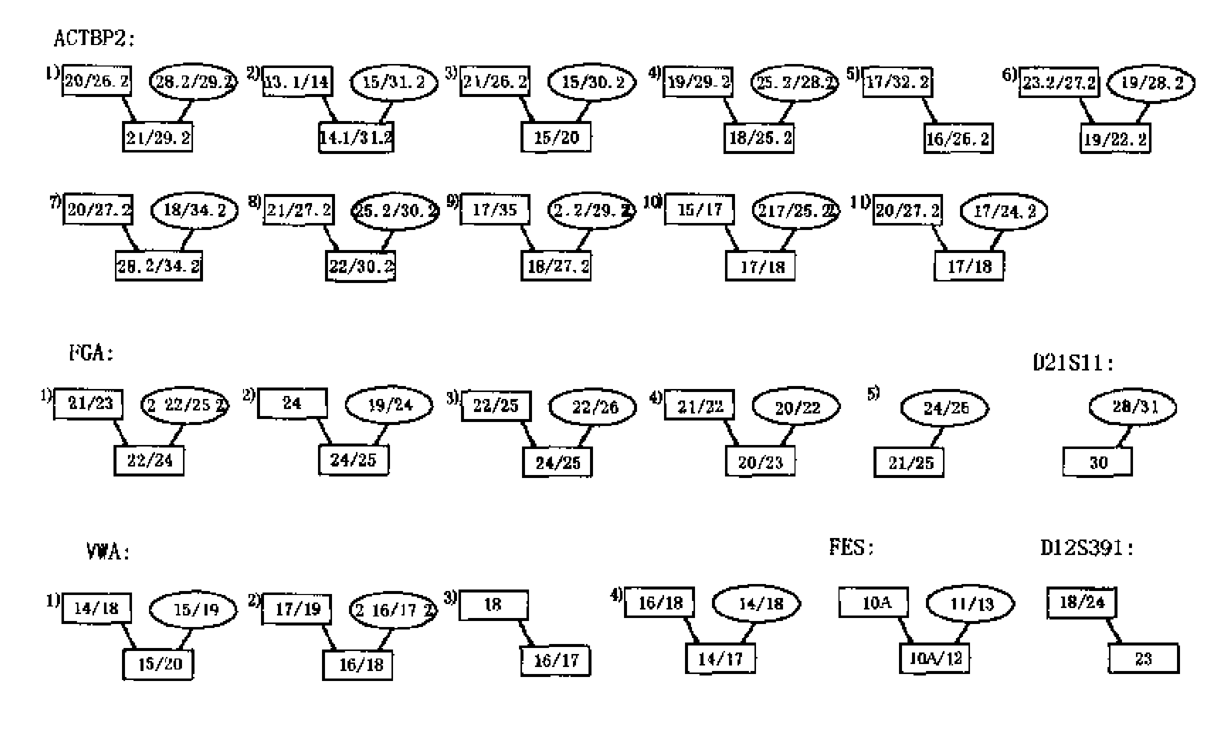


图 12-7 从 10844 次亲代/子代等位基因传递中发现的 23 个突变

突变率与连续重复单位数的平均值成正相关性，与基因座的结构、长度有关。平均值

越大，突变率越高。一般重复单位数小于10的基因座不易发生突变，而单一重复单位连续重复10次以上的等位基因易发生突变。等位基因内部不含有完全重复单位的等位基因比含有不完全重复单位的等位基因易发生突变。例如在图12-8中，CD4，F13B和TH01基因座的同一重复单位平均值小于10，因此没有观察到突变；D21S11基因座等位基因30易发生突变，但等位基因32.2不易发生突变，因为等位基因32.2含有TA的插入。同样地，ACTBP2基因座突变发生在较小等位基因区，而在较大等位基因区反而低，这也是因为较大等位基因的内部都含有AAAAAG序列，AAAG重复单位被它隔断，这些较大等位基因实际的同一重复单位平均值反而比小等位基因区的小。重复单位侧翼或附近的不规律的重复单位存在对突变没有限制作用，例如图中的FGA基因座，旁侧的不完全重复单位并不减少同一重复单位几何平均值，不能减缓突变的发生。

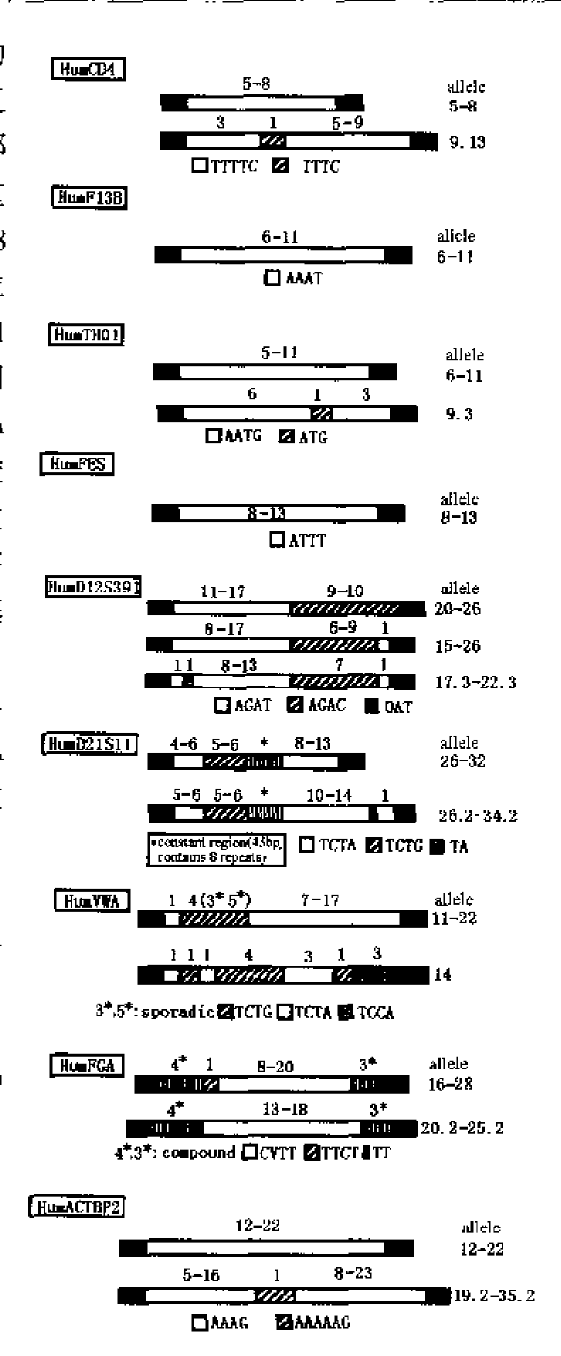


图12-8　几个基因座等位基因结构

在同一系统中，突变率与等位基因重复数有较大关系，等位基因长度越长，突变发生的可能性越大。

在男性和女性之间，突变的发生有差异。父方等位基因的突变比母方的更频繁，Brinkmann（1998）发现男性与女性的突变比例为17:3，男性精子分化经历的细胞分裂数比卵子的高10倍，Y-染色体的取代积累比常染色体快2倍。如果突变率与细胞分裂成正比，据估计，卵原细胞在减数分裂之前经历了~22次分裂，形成卵细胞。精原细胞靠继续的有丝分裂再生，在形成精子之前，它们经历了更多的分裂。对于一个29岁的男子，在进入减数分裂前，大约经历了350次分裂。所以对29岁的男性，突变率比女性的高16倍。而且，随着父亲年龄的增长，单个重复单位的突变发生率增加。而在女性中则没有这种年龄效应。

3. 部分常用STR基因座的突变率

STR基因座的平均突变率在0.1%以下，也就说必须调查近1 000个亲代-子代间的等位基因传递，才能在一些基因座观察到一个突变。Brinkman（1998）等调查了10 844个父母-子在9个STR基因座的等位基因传递，观察到23个突变，在TH01，F13B和CD4等3个基因座没有发现有突变。Sajantila（1999）研究29 640个父-子-母在5个STR基因

座和 4 个小卫星基因座的等位基因传递，只观察到 18 个突变，其中在 3 个 STR 基因座 D3S1359，VWA 和 TH01 中观察到 11 次，TPOX 和 FES/FPS2 个基因座中未发现突变。表 12-6 为美国血库协会调查的 CODIS 系统 13 个 STR 基因座的突变率。表 12-7、表 12-8 和表 12-9 是一些文献报道的 STR 基因座突变率。大多数突变率为 0.1%～0.5%。其中 CSF1PO，TH01，TPOX，D5S818 和 D8S1179 基因座的突变率最低，D21S11，FGA，D7S820，D16S539 和 D18S51 基因座的突变率偏高。

表 12-6 CODIS 系统 13 个基因座的突变率（美国血库协会 1999）

基因座	观察到的突变率	
	母方	父方
CSF1PO	14/47843（0.03）	311/243124（0.13）
FGA	7/8253（0.01）	555/189973（0.29）
TH01	5/42100（0.01）	12/74426（0.02）
TPOX	2/28766（0.01）	10/45274（0.02）
vWA	20/58829（0.03）	851/250131（0.34）
D3S1358	0/4889（<0.02）	9/8029（0.11）
D5S818	22/60907（0.04）	194/138033（0.150
D7S820	14/50827（0.03）	193/1318880（0.15）
D8S1170	5/6672（0.07）	29/10952（0.260
D13S317	33/59500（0.06）	106/69598（0.15）
D16S539	12/42648（0.03）	40/48760（0.08）
D18S51	8/8827（0.09）	29/9567（0.30）
D21S11	12/6754（0.18）	17/6980（0.24）

表 12-7 9 个 STR 基因座的突变率（Brinkmann，1998）

基因座	有效长度[a]	减数分裂数[b]	突变数	突变率（%）	杂合度[c]
CD4	5.66	969	0	0.000	0.67
F13B	8.74	1033	0	0.000	0.67
TH01	7.78	2008	0	0.000	0.78
FES	10.79	850	1	0.117	0.65
vWA	10.8	2013	4	0.199	0.81
D12S391	11.47	562	1	0.178	0.87
D21S11	11.0	557	1	0.180	0.86
FGA	13.84	1246	5	0.401	0.96
ACTBP2	15.21	1608	11	0.684	0.93

a：同一连续重复单位数的几何平均值；b：父、母方等位基因的传递总数；c：德国人群数据

表 12－8　Henke（1999）报道的父、母方突变率

基因座	突变率	
	父方	母方
CSF1PO	0/237	0/165
D13S317	0/258	0/178
D18S51	0/286	2/205
D21S51	1/267	3/189
D3S1358	0/257	0/176
D5S818	0/258	0/178
D7S820	0256	2/176
D8S1179	0/213	0/149
FGA	0/307	3/218
ACTBP2	0/402	5/315
TH01	0/394	0/301
TPOX	0/240	0/167
vWA	1/258	0/178

表 12－9　其他部分研究人员报道的突变率

基因座	突变率	出处
TH01	1/5918	Sajantila et al 1999
TPOX	0/2457	Sajantila et al 1999
vWA	5/3684	Sajantila et al 1999
D3S1358	1/390（父方），1/390（母方）	Szibor et al，1998
	1/130	Mornhinweg，1998
D13S317	1/217	Lins，1998
D16S539	0/300	

4. 突变对亲子鉴定的影响

突变对亲子鉴定影响很大。因为孩子与假定父亲间关系确定是基于等位基因按孟德尔遗传方式传递的假设前提。如果不注意 STR 基因座突变，可能导致误排生父；对于群体进化研究而言，群体统计和群体历史必须减去突变过程才能正确研究群体遗传问题。一些群体性灾难事件中残骸的身份确定与一些尸源鉴定，要与嫌疑家庭人员进行比对，突变将是一个值得注意的重要问题，因为如有突变存在，将会出现不完全的匹配。

12.1.8 用于法医 DNA 分型的理想 STR 基因座特点

目前在人类基因组中已发现 8 000 多个 STR 基因座，但不是所有的 STR 基因座都适用于法医 DNA 分析应用。从实用价值考虑，STR 基因座应具备以下特点：

1. 等位基因片段长度在 95～500bp 范围，最好在 400bp 以下；
2. 具有较高的杂合度及个人识别能力，一般杂合度在 0.70 以上，DP > 0.90；
3. 具有规律的四核苷酸或五核苷酸重复序列，阴影带出现率较低，较少或没有插人

等位基因；

4. 等位基因数为 10～12 个，等位基因易于区分；

5. 与其他基因座复合扩增时，重复性好；

6. 突变率低，不超过 0.2%；

7. 具有高度种属特异性。

8. 复合扩增体系中的各 STR 基因座分布在不同的染色体上，基因座之间没有连锁，可利用乘法原则计算偶合率。

12.1.9 **阴影带或峰**

阴影带（stutter band）或峰（stutter peak），指的是等位基因主带前（短一个重复的位置，图 12－9）的一些强度较浅的带，经测序证实其长度比等位基因短一个重复序列，阴影带是在复制过程中 Taq DNA 聚合酶的滑动造成的（图 12－10）。复制滑动在二核苷酸 STR 基因座较为多见；四、五核苷酸重复基因座则相对少见。在复合扩增中，由于扩增反应条件不是每个基因座的最佳条件，容易出现阴影带。

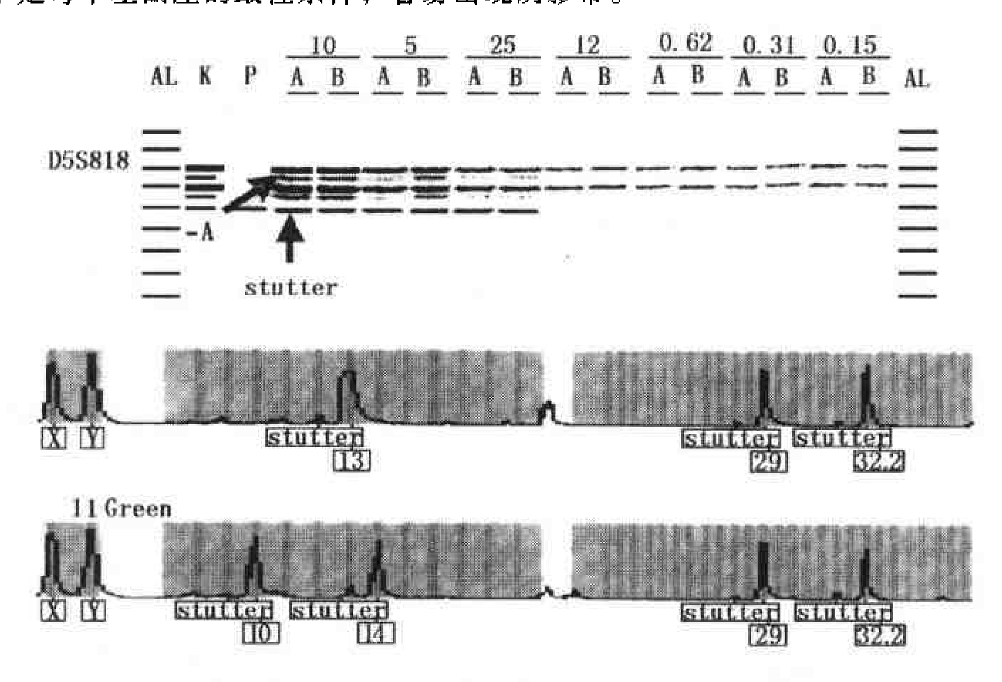


图 12－9 stutter 峰图中所标有 stutter 即为 stutter 峰

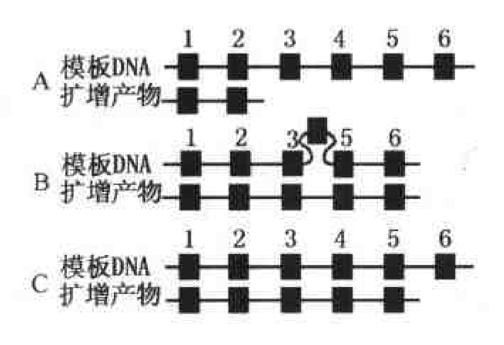


图 12－10 产生 Stutter 的滑动机制

阴影带发生频率不仅与复合体系中的基因座有关，而且同一基因座内不同等位基因也不同，一般随等位基因长度增加而增加。阴影带峰高与等位基因峰高的百分比叫“stutter”百分比，用来表示阴影带的强度，它有以下一些特点：（1）“stutter”百分比一般在10%以下，在同一基因座内随等位基因长度增加而增加（图12－11）；（2）“stutter”百分比重复性很好，它们平均偏差在±2.5%范围内；（3）阴影带不仅取决于基因座，也取决于所用的聚合酶和PCR反应条件；（4）阴影带随重复单位的长度增加而减少，五聚核苷酸的阴影带＜四核苷酸＜三核苷酸＜二核苷酸；（5）含有不完整重复序列的等位基因较少产生阴影带；（6）在150～4500RFU（相对荧光素单位）之间“stutter”百分比不随加样DNA量变化而变化。对于范围外（荧光检测仪器设定的最高峰的阈值）的等位基因峰，“stutter”百分比可能不正常的高，在电泳时减少PCR产物上样量，可以增加定量测定的准确性。

阴影带的存在会增加混合样品DNA结果解释的复杂性，比如当混合样品中某一个成分的含量比较低时，很难区分开“stutter”峰与混合样品中较少组分的等位基因，使结果解释变得复杂，因此必须予于注意。

减少阴影带产生有以下三种方法：

（1）采用较长重复单位的STR基因座。目前已发现一些五核苷酸重复序列的基因座，它们的平均Stutter百分比低于1%。

（2）含有不完全重复单位的等位基因产生的阴影带少。例如，TH01基因座的核心序列为AATG，等位基因9.3在重复区中间含有一个ATG序列，当核心序列被打断时，产生的阴影带比相近长度含有完整核心序列的等位基因低，VWA的情况也证实这点。

（3）采用延伸速度快的DNA聚合酶。阴影带的形成与DNA聚合酶复制模板的速度有关，如果聚合酶复制速度慢，则相应的阴影带量增加。有人提出如果将来热稳定DNA聚合酶的速度比现在Taq DNA聚合酶的50－60碱基/秒速度要快，那么阴影带就会减少。

12.1.10 3′端的额外A核苷酸添加

Taq DNA聚合酶具有类似DNA末端转移酶（TdT）的功能，可在新合成双链产物的3′末端加上一个非模板依赖的碱基，使PCR产物3′－端经常有额外的碱基，通常称之为非模板添加物。理论上四种碱基均可被添加到3′端，但Taq DNA酶对dATP的聚合能力远远高于其它三种dNTP，PCR产物3′端的非模板添加以腺嘌呤（adenine）为主，称为A添加，这种非模板添加导致PCR产物比真正目的序列多1bp。在PCR扩增中最后一步60℃或72℃保温有利于A的添加，而且加A的程度取决于模板链的序列与反向引物的5′端序列。有人发现引物5′端为G，利于聚合酶的A添加。因此引物序列不同，每个基因座的腺化比例稍有差异。

如果PCR产物部分腺化，而不是全部加A，使每个等位基因有大小相差1bp的二个片段（“n”和“n＋1”带，n为全长等位基因，n＋1带为等位基因＋1，或表达为＋A/－A），表现为图12－12所示的肩膀峰（shoulder peak）或为裂开的峰（split peak）。“n”和“n＋1”带的存在增加分型复杂性，无法区分是差1bp的二个真正等位基因还是不完全腺化造成的“n”和“n＋1”带，比如，TH01基因座的等位基因10没有腺化的片段长度与等位基因9.3腺化的长度相同。因此要求所有PCR产物要么加A，或要么都不加A。表12－10列出了＋A与－A的方法。

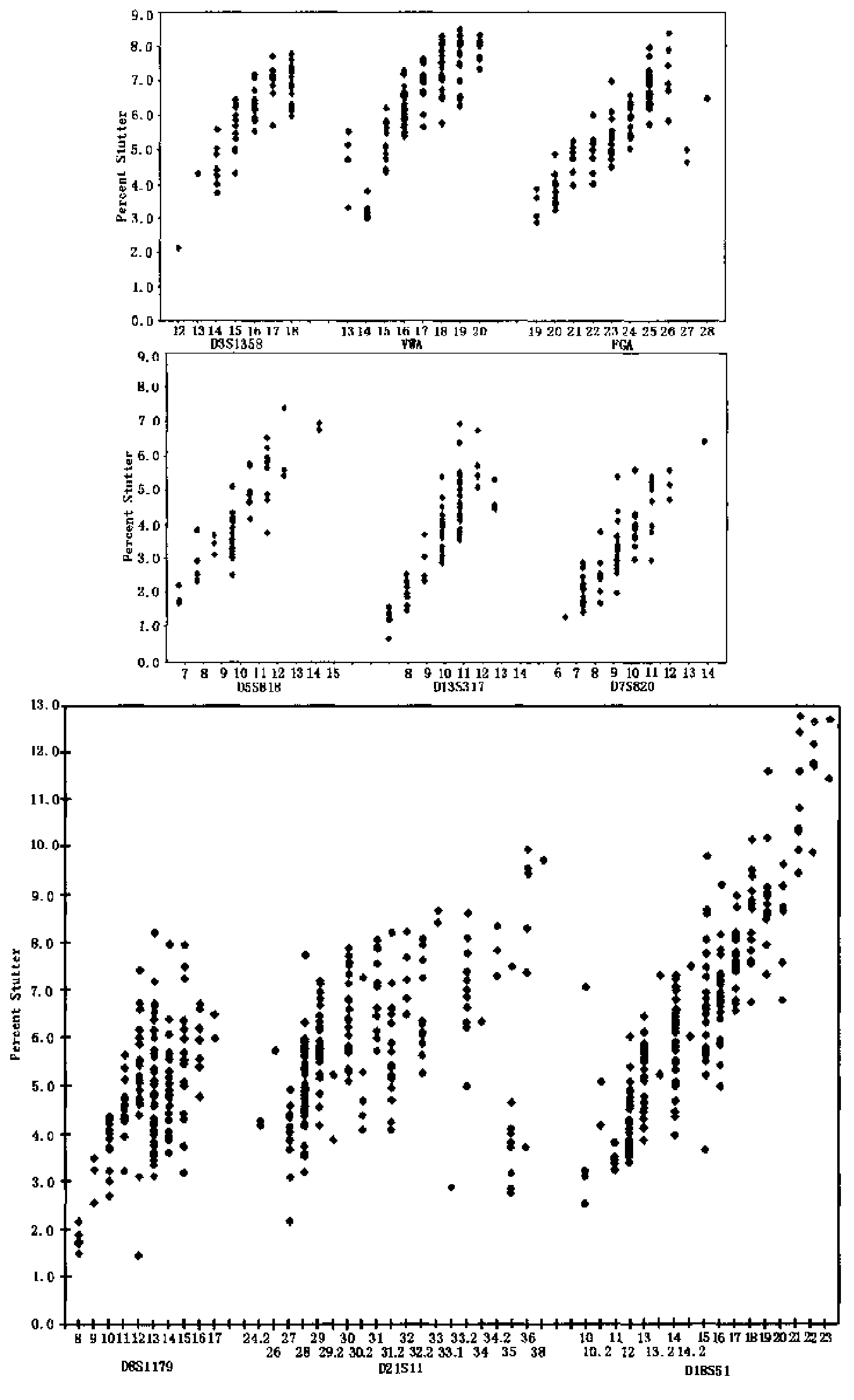


图 12-11　一些基因座的 stutter 百分比

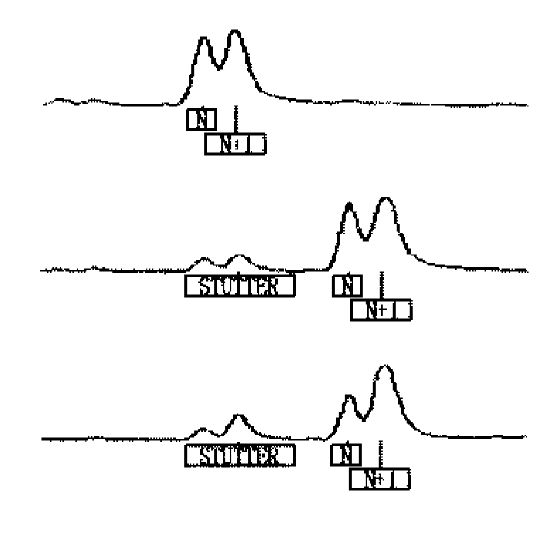


图 12－12 ＋A 与－A 带

表 12－10 末端加 A 与不加 A 的方法

方法	结果
产物全部加 A：	
在 60℃或 72℃延伸 30～45min	使所有产物全部腺化
反向链引物 5′端加— GTTCTT 序列	整链引物 3′端完全加 A
产物不加 A：	
反向链引物加入一个内切酶切点	内切酶消化切去添加的 A
酶除去添加的 A	利用 pfu 或 T4 聚合酶的外切酶活性除去添加的 A
采用没有末端转移酶的修饰酶	聚合酶没有 3′端加 A 功能

在实际使用中，采取使所有 PCR 产物的 3′端进行非模板添加。保证最大可能 A 添加的方法是：一是选择有利于 A 添加的引物序列；二是扩增最后一个循环是 60℃下保温 45min。

12.2 STR 基因座检验

12.2.1 多 STR 基因座复合扩增

1. 复合扩增的特点

STR 基因座由于串联重复序列只有 2～7bp，等位基因片段比较小，等位基因数目相对有限，片段长度范围小，较容易扩增，各基因座之间扩增条件基本相似，因此几个不同的基因座可以安排在同一个反应管中进行扩增，这种把二个或二个以上的基因座放在同一个反应管中同时扩增，叫复合扩增。单个 STR 基因座等位基因数少，信息量有限，实际案件鉴定常需联合检测 8～12 个基因座，才能达到类似 DNA 指纹的同一认定水平；多个 STR 基因座复合扩增，明显提高了单次检测信息量，一次检测可以达到个人同一认定，但灵敏度比单基因座扩增低。例如四基因座复合扩增体系，其灵敏度为 100pg DNA，低于 100pg 基因组 DNA 复合扩增时会出现一些等位基因丢失现象。

不论 STR 的单基因座扩增还是多基因座复合扩增，PCR 反应体系仍由 PCR 反应缓冲

液、dNTPs、模板 DNA、引物、DNA 聚合酶及 Mg^{2+} 成分组成，只是复合扩增中引物的数量相对多些。在确定复合扩增反应热循环参数时，扩增条件并不是每个基因座所特定的、最适的。因此，在实验设计时要考虑到各基因座间扩增效率的均衡。

2. 复合体系基因座数目的确定

多少个基因座可以在一个复合体系中扩增，取决于扩增产物检测系统。目前常用检测系统有银染和荧光染料标记自动检测系统。

（1）银染系统　银染法只根据各等位基因的片段大小区分，要求组成复合扩增的各个基因座等位基因片段长度范围没有重叠。复合体系中基因座数目不宜多，一般为 3～4 个。

（2）多色荧光标记检测系统　该检测系统可以根据等位基因片段所带的荧光颜色不同加以鉴别，即使等位基因片段长度有重叠的基因座，也可采取不同的荧光染料标记，区分不同基因座的等位基因。因此这种复合体系可以有较多的基因座，如 7 个基因座、8 或 10 个基因座复合扩增，甚至开发有 16 个基因座的复合扩增系统，明显提高了检测的信息量。

12.2.2 扩增参数对 STR 复合扩增的影响

对于某一特定 STR 基因座，PCR 反应参数是特定的，也是最适的，当二个或二个以上基因座在同一反应体系中进行复合扩增时，要兼顾各个基因座的特点，合理调整扩增体系成份、热循环参数等因素，才能提高复合体系扩增效率。

1. PCR 缓冲液

Taq DNA 聚合酶的标准反应缓冲液成分为 10mmol/L Tris - HCl（pH8.3），50mmol/L KCl，1.5mmol/L $MgCl_2$，1%TritonX100 或 0.001%白明胶。1% Triton - X100 和 0.01%白明胶都具有缓解蛋白质对 Taq DNA 聚合酶的抑制作用，不同厂家 PCR 缓冲液没有明显差异。

PCR 反应缓冲液浓度改变会影响 PCR 扩增，例如图 12 - 13，当标准缓冲液浓度减少到一半时，复合扩增 PCR 产物会减少，不同基因座受影响程度不同；当浓度增加到两倍时，背景信号也明显增加。

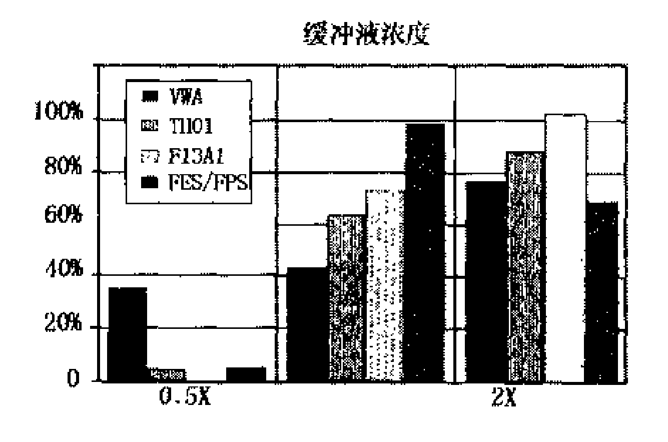


图 12 - 13　不同 PCR 缓冲液浓度对四重复合（vWA、F13A1、FES/FPS、TH01）扩增效率的影响

2. 引物浓度

（1）复合体系中某基因座引物浓度发生改变，该基因座扩增产物信号强度会发生变

化，但基本不影响复合体系中其他基因座，信号没有明显变化。但有个别基因座引物浓度变化，如 vWA，可能会引起其他基因座信号强度降低（图 12－14）。

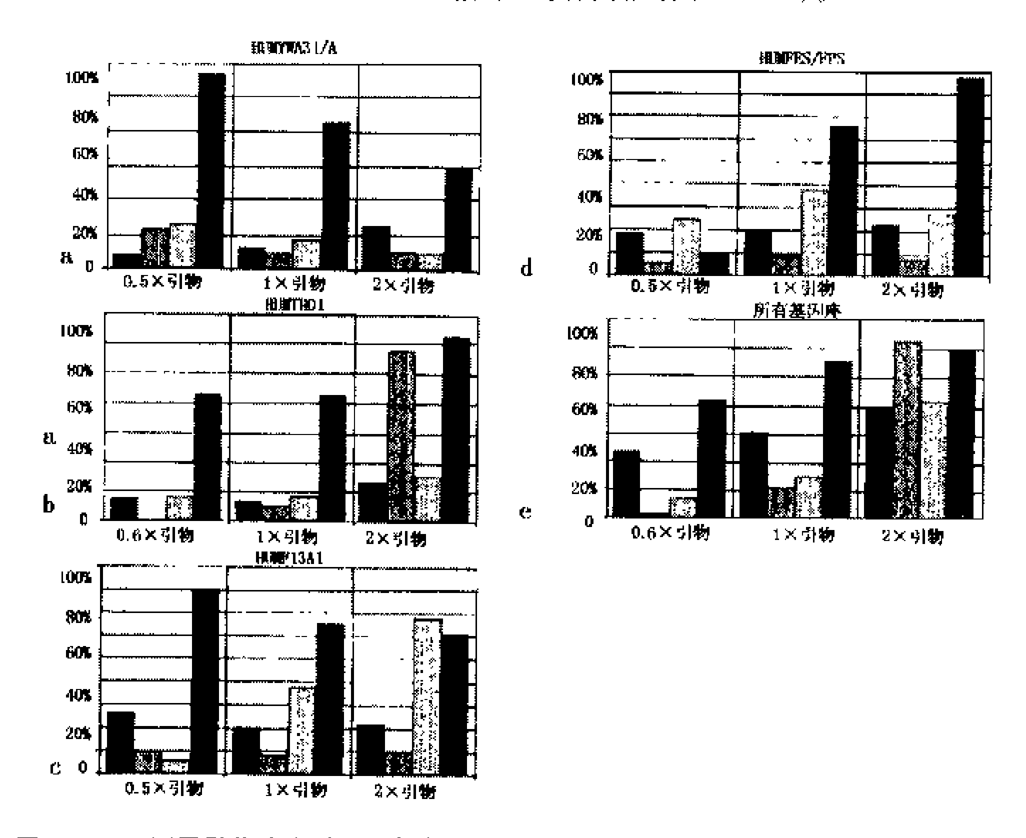


图 12－14　不同引物浓度对四重复合（vWA、F13A1、FES/FPS、TH01）扩增效率的影响

a－d：单个基因座引物变化；e：所有引物浓度变化

（2）复合体系中所有基因座引物浓度同时发生变化，所有基因座扩增均受到影响。例如引物浓度减半，所有基因座信号均降低；引物浓度增加一倍，所有基因座信号增加，引物二聚体产生的杂带出现频率也增加。

（3）从复合体系中减去任一基因座引物，对其他基因座扩增没有明显影响。

3.dNTP

在 PCR 反应中，dNTPs 与 Mg^{2+} 定量结合，最适 dNTPs 浓度与 Mg^{2+} 浓度直接相关。在标准反应体系中，dNTPs 总浓度为 800 μmol/L（每一个 dNTP 浓度为 200μmol/L）。dNTPs 浓度减少，聚合酶链反应底物不足，所有扩增产物减少，减少程度随基因座不同而不同，dNTPs 浓度超过反应最佳水平时，所有基因座扩增效率明显降低，这可能是由于游离的 Mg^{2+} 不断减少，聚合酶活性降低了（图 12－15a）。

由于 Taq 酶具有非模板的 PCR 产物 3′端一个碱基添加功能，PCR 产物中存在正常长度和 3′端多一个额外碱基的产物（即前面介绍的 n 与 n＋1 带）共存，dNTPs 浓度增加时，游离的 Mg^{2+} 减少，Taq 酶的 3′端非模板添加活力也降低，n＋1 带的比例下降。

4.Taq 聚合酶浓度

50μl 反应体系中 Taq 酶用量为 1.0～2.5u，当反应体系中 Taq 酶量大于 2.5u，非特异带扩增就会增加；少于 1.0u，PCR 特异产物较少或甚至没有（图 12－15b）。

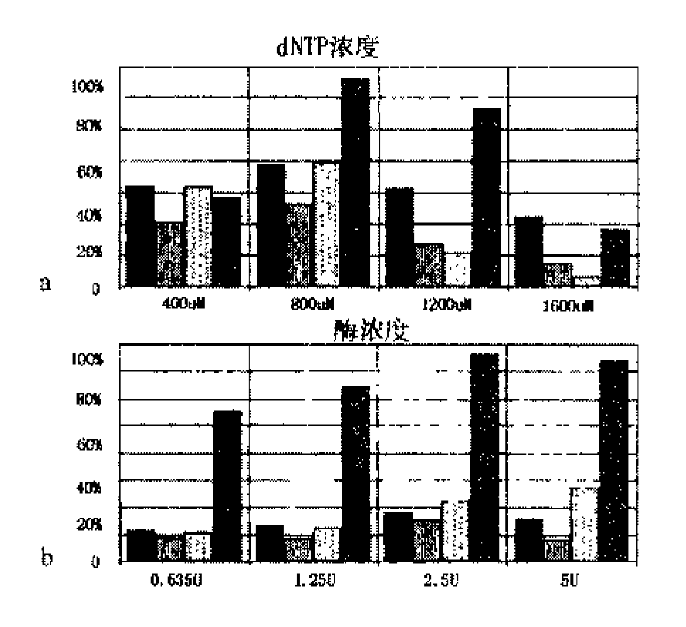


图 12-15 dNTP 浓度和酶浓度对四重复合（vWA、F13A1、FES/FPS、TH01）扩增效率的影响

a：dNTP 浓度；b：酶浓度

5. 模板 DNA 浓度

模板 DNA 浓度减少，扩增产量降低，不同基因座减少程度不同（图 12-16）。模板 DNA 拷贝数很低时，可引起同一基因座二个等位基因的不平衡扩增，表 12-11 是杂合子个体在模板 DNA 100pg~1ng 范围内二个等位基因强度信号比。当模板 DNA 为 100pg 时，不平衡扩增造成荧光信号强度比可高达 56%，因此在对微量样品结果分析时，应当注意等位基因间的不平衡扩增。模板 DNA 低于 100pg DNA 时不平衡现象更加严重，当模板 DNA 低于 50pg 或更少，其中较大一个等位基因得不到扩增，易造成假纯合子现象；而且模板 DNA 量少，不同基因座之间扩增效率也不平衡，扩增片段长度较大的基因座的扩增产物明显减少。

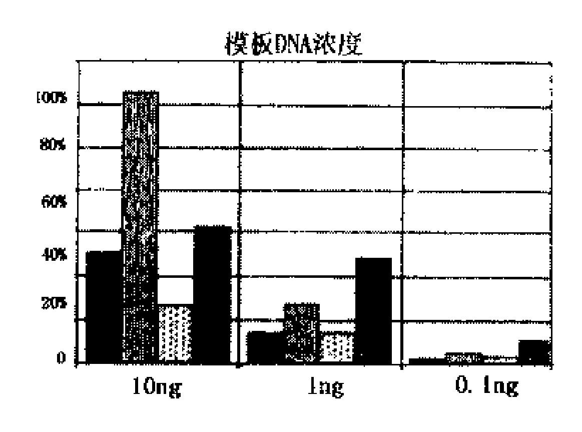


图 12-16 模板 DNA 浓度对四重复合（vWA、F13A1、FES/FPS、TH01）扩增效率的影响

表 12－11　模板 DNA 量对杂合子个体复合扩增效率的影响

模板 DNA 量	杂合子等位基因信号强度差异大于 20% 的样品百分比			
	HUMvWA31/A	HUMTH01	HUMF13A1	HUMFES/FPS
100pg	41%（46%）	46%（51%）	39%（56%）	34%（41%）
250pg	25%（30%）	39%（44%）	17%（40%）	23%（60%）
500pg	18%（28%）	18%（41%）	5%（21%）	5%（30%）
750pg	27%（28%）	47%（28%）	5%（21%）	10%（28%）
1ng	18%（28%）	6%（28%）	0%（15%）	3%（21%）

括号内的数字为观察到的等位基因信号最大值。

6. 扩增循环数

在 PCR 反应进入平台期前，PCR 产物总产量与 PCR 循环数直接相关。进入平台期，PCR 产量不再明显增加，PCR 循环数一般为 28～35 个。不同基因座进入平台期的时间不同，所以在同一个循环数下，各基因座的产量有所差异（图 12－17）。

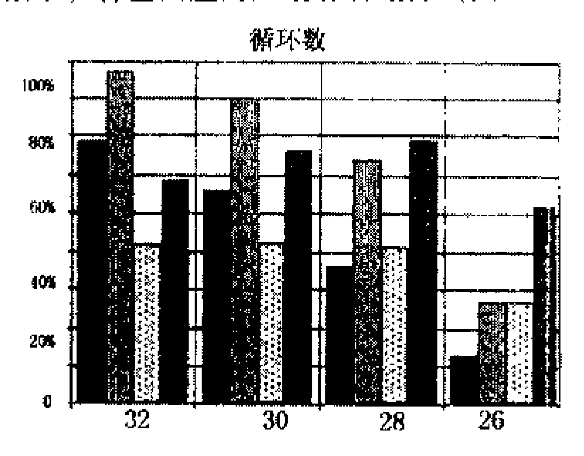


图 12－17　扩增循环数对复合扩增效率的影响

7. 变性温度

PCR 反应的变性温度一般为 93～95℃，在此范围内基本没有差异，当变性温度为 97℃时，PCR 扩增效率明显降低，因为温度升高，使 Taq 酶半衰期变短，如 95℃半衰期为 40min，而 97℃时则为 5min。不同基因座间受影响程度有所差异（如图 12－18a）。

8. 退火温度

每一个引物碱基序列不同，GC 含量不同，最适退火温度也不同，退火温度变化引起不同基因座扩增效率改变。退火温度提高，信号减弱；退火温度低，信号强，同时背景信号强；退火温度升高或降低，对复合体系中任一基因座的影响结果与单一扩增情况相同（图 12－18b）。

9. 非特异扩增产物

在复合体系中，存在多对引物，易发生引物与引物之间的结合，形成引物二聚体（primer－dimer），产生一些低分子量（55～75bp）非特异性产物。较少量的引物二聚体存在，并不影响扩增产物产量，但是较多的引物二聚体会导致扩增产量降低，甚至严重时检

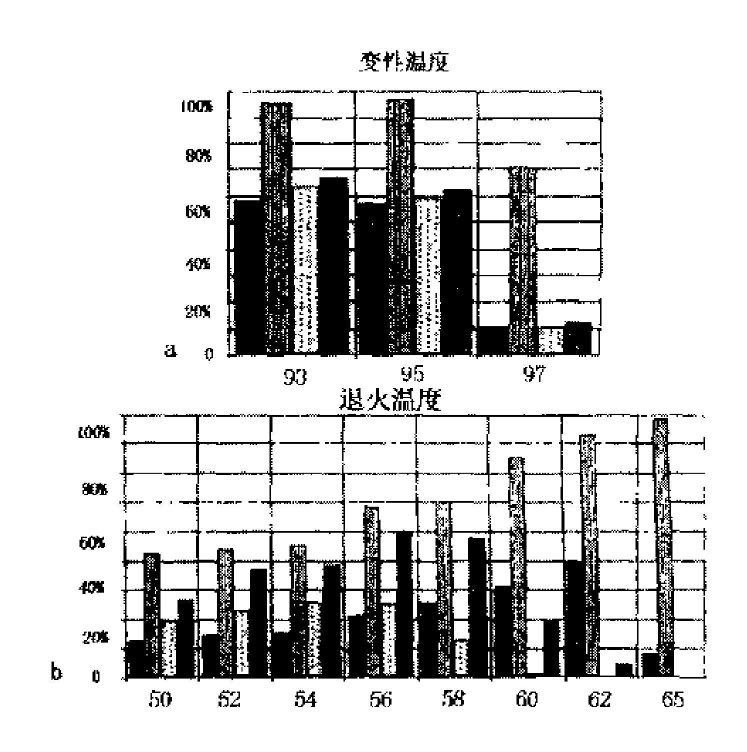


图 12-18 变性温度和退火温度对四重复合（vWA、F13A1、FES/FPS、TH01）扩增效率的影响

a：变性温度；b：退火温度

测不到 PCR 产物。

在 PCR 反应前，引物二聚体的形成随引物总浓度增加而增加。另外，引物二聚体形成与 PCR 反应循环前试剂温度有关，当 PCR 反应液温度升高到 70℃，引物二聚体大量形成。因为较高温度增加 Taq 酶活力，因此在准备 PCR 反应时最好在低温下进行。

10. 离子强度

改变离子浓度也会影响 PCR 反应，如反应体系中 NaCl（或 KCl）终浓度大于50mmol/L 时（表 12-12），Taq 酶活力受到抑制，PCR 扩增受阻。

表 12-12 离子强度对扩增效率的影响

基因座	等位基因信号降低百分比			
	NaCl 浓度			
	125mmol/L	100mmol/L	75mmol/L	50mmol/L
VWA	NS	50%	0%	0%
TH01	NS	50%	25%	0%
F13A1	NS	NS	NS	0%
FES/FPS	NS	NS	NS	0%

NS 表示没有信号检测到

11. pH 值

在酸或碱存在下，扩增效率取决于 PCR 缓冲液能力。当 5mmol/L HCl 或 7.5mmol/L

NaOH 存在下，PCR 扩增完全受到抑制，在较低 HCl 或 NaOH 下，部分扩增产物信号降低（表 12 - 13）。

表 12 - 13　pH 值对扩增效率的影响

基因座	等位基因信号降低百分比									
	HCl 最终浓度（mmol/L）					NaOH 最终浓度（mmol/L）				
	7.5	5	2.5	1	0.5	0.5	1	2.5	5	7.5
VWA	NS	NS	60%	10%	10%	0%	0%	10%	20%	NS
TH01	NS	NS	50%	0%	0%	0%	0%	0%	10%	NS
F13A1	NS	NS	80%	10%	0%	0%	0%	10%	30%	NS
FES/FPS	NS	N	90%	20%	10%	0%	0%	10%	50%	NS

NS 表示没有信号检测到

12. 引物之间配比

为获得尽可能平衡的信号强度，引物浓度需调整。确定最适引物浓度还取决于 PCR 产物检测方法。当用荧光检测系统时，涉及二个主要因素：（1）不同荧光染料的检测灵敏度不同；（2）引物标记效率和纯化，对于新合成或标记纯化的一批引物，首先必须确保它在复合体系中的特异性。

12.2.3 复合扩增 STR 基因座的要求与标准

对于法医 DNA 分析应用的复合扩增体系的 STR 基因座，除了满足一般应用要求外，还有如下要求：

1. 复合体系中最好含有 1 个性别基因座（如 Amelogenin）；

2. 所有基因座分布在不同染色体上，保证基因座之间没有连锁，以利系统可以用乘法原则计算偶合率。

3. 包含几个低分子量的基因座（扩增产物片段小于 150bp）。

4. 采用特异性好，酶活性高的 Taq DNA 聚合酶，如 Gold Taq DNA 聚合酶。

5. 各基因座应有相应的等位基因分型标准物，便于分型。

6. 对于银染检测系统，各基因座的等位基因片段长度不能重叠，复合体系不宜超过 4 个 STR 基因座。

12.2.4 复合扩增类型

自 1992 年初 Edward 等正式报导了 3 个基因座 STR 复合扩增以后，陆续有作者相继报道了新的复合扩增体系。按复合扩增 STR 基因座的多少分为：双基因座复合扩增，三基因座复合扩增，四基因座复合扩增，六基因座复合扩增，七基因座复合扩增，以及八、九、十、十三、十五、十六等基因座复合扩增等。复合体系中基因座数越多，体系的个体识别率越高，鉴定能力越强，表 12 - 14 列出了部分商业试剂盒。

表 12-14 部分试剂盒

试剂盒名	生产厂家	包含的基因座	识别率
TH01、TPOX、CSF1PO monoplexes（银染系列）	Promega	TH01、TPOX、CSF1PO	1:410
AmpFlSTR^R Blue	PEApplied Biosystems	D3S1358、VWA、FGA	1:5000
AmpFlSTR^R Green I	PEApplied Biosystems	Amelogenin，TH01，TPOX，CSF1PO	1:410
CTTv	Promega	CSF1PO，TPOX，TH01，VWA（vWF）	1:6600
FFFL	Promega	F13A1，FES/FPS，F13B，LPL	1:1500
GammaSTR	Promega	D16S539，D13S317，D7S820，D5S818	$1:1.8\times10^4$
Powerplex^{TM}（1.1 和 1.2）	Promega	CSF1PO，TPOX，TH01，VWA，D16S539，D13S317，D7S820，D5S818	$1:1.2\times10^8$
AmpFlSTR^R profiler	PEApplied Biosystems	D3S1358，VWA，FGA，Amelogenin，D8S1179，D21S11，D18S51，D5S818，D13S317，D7S820	$1:9.6\times10^{10}$
AmpFlSTR^R Cofiler	PEApplied Biosystems	D3S1358，D16S539，Amelogenin，TH01，TPOX，CSF1PO，D7S820	$1:8.4\times10^5$
AmpFlSTR SGM plus	PEApplied Biosystems	D3S1358，VWA，D16S539，D2S1338，Amelogenin，D8S1179，D21S11，D18S51，D19S433，TH01，FGA	$1:3.3\times10^{12}$
Powerplex^{TM}2.1	Promega	D3S1358，TH01，D21S11，D18S51，VWA，D8S1179，TPOX，FGA，PentaE	$1:8.5\times10^{10}$
Powerplex^{TM}16	Promega	CSF1PO，FGA，TPOX，TH01，VWA，D3S1358，D5S818，D7S820，D8S1179，D13S317，D16S539，D18S51，D21S11，Penta D，Penta E，Amelogenin	$1:1.8\times10^{17}$

需要注意的是，即使是同一个基因座，但不同试剂盒中的引物序列可能不同，等位基因片段长度有所不同，例如 vWA 基因座，Promega 公司的 PowerplexTM1.1 试剂盒比 PEApplied Biosystems 的相应等位基因大 29bp、其他基因座也有类似现象（表 12-15），而且还会出现个别个体的基因型不一致。

表 12-15 不同试剂盒相同基因座的扩增等位基因片段差异

基因座	染色体定位	核心序列	
CSF1PO（6~15）	5qter	AGAT	Promega 公司的 Powerplex^{TM}1.1 试剂盒比 PE Applied Biosystems 的大 11bp
FGA（11~51.2）	4qter	CTTT	Powerplex^{TM}2.1 试剂盒比 PE Applied Biosystems 的大 112 bp
TH01	11pter	AATG	Powerplex^{TM}2.1 试剂盒比 PE Applied Biosystems 的大 11 bp，Powerplex^{TM}1.1 与 Powerplex^{TM}2.1 之间相差 19 bp

续表

TPOX	2pter	AATG	Powerplex^{TM}2.1 试剂盒比 PE Applied Biosystems 的大 7 bp，Powerplex^{TM}1.1 与 Powerplex^{TM}2.1 之间相差 38 bp
VWA	12pter	TCTA，TCTG	Powerplex^{TM}2.1 试剂盒比 PE Applied Biosystems 的小 29 bp，Powerplex^{TM}1.1 与 Powerplex^{TM}2.1 之间没有差异
D3S1358	3pter	ACA，AGAC	PE Applied Biosystems 的 multiplex Profiler，Profiler Plus，COfiler 及 SGM Plus 试剂盒的等位基因大小一致，它们比 Powerplex^{TM}2.1 的小 2 bp
D5S818 (7～16)	5qter	AGAT	Powerplex^{TM}1.1 试剂盒比 PE Applied Biosystems 的小 15 bp
D7S820	7qter	GATA	Powerplex^{TM}1.1 试剂盒比 PE Applied Biosystems 的小 42 bp
D8S1179	8	TCTA，TCTG	Powerplex^{TM}2.1 试剂盒比 PE Applied Biosystems 的大 80 bp
D13S317	13qter	GATA	Powerplex^{TM}1.1 试剂盒比 PE Applied Biosystems 的小 36 bp
D16S539	16qter	AGAT	Promega 公司的试剂盒比 PE Applied Biosystems 的小 31 bp
D18S51	18qter	AGAA	Powerplex^{TM}2.1 试剂盒比 PE Applied Biosystems 的大 22 bp
D21S11	21	TCTA，TCTG	Powerplex^{TM}2.1 试剂盒比 PE Applied Biosystems 的大 17 bp

12.2.5 常用的 STR 复合扩增基本技术

STR 基因座扩增的原理和小卫星的 VNTR 基因座基本相同，只是 PCR 反应体系中组成成分含量不同，模板 DNA 用量减少，扩增退火、延伸时间较小卫星 VNTR 短。

为了方便操作、保证重复性，预先配置引物混合液（引物－mix）。当然也可以采用与小卫星 VNTR 一样的各组分各自单独加样的方式。以下所介绍的只是一般性的操作程序，反应体系各组分的含量要根据具体检验的基因座进行调整，达到最佳反应体系组成以及扩增热循环参数。对于使用商业化的试剂盒应严格按照说明书操作。

1. 溶液

（1）10×STR 缓冲液：100mmol/L Tris－HCl，pH 8.3，500mmol/L KCl，15mmol/L $MgCl_2$，10% Triton X－100，灭菌，保存在－20℃。

（2）引物混合液：0.5mmol/L 各 dNTP，500nmol/L 各引物。

（3）PCR 体系：10mmol/L Tris－HCl（pH 8.3），50mmol/L KCl，1.5mmol/L $MgCl_2$，1% Triton X－100，0.2mmol/L 各 dNTP，200nmol/L 各引物，0.025u/μl Taq DNA 聚合酶。引物浓度应根据具体基因座而定。

2. 操作方法

（1）计算所需要试剂的量：每一反应管所需的量×（所要分析的样品数 ＋3）。多出的 3 份主要是：一份用于阳性对照，一份用于阴性对照，另一份用于补偿加样过程中的损失；

（2）在室温下融化引物；

（3）标记每一个反应管；

(4) 将 PCR 反应管放在冰水浴中架子上 10s;

(5) 加入所需的无菌水(27.5μl/反应)、10×STR 缓冲液、引物混合液(20μl/反应);

(6) 加入所需 Taq DNA 聚合酶量(0.25μl 的 5u/μl Taq 聚合酶),混匀,即配制成 PCR 混合液;

(7) 每一个反应管中加入 48μl 的 PCR 混合液;

(8) 加 2μl DNA 提取液,总体积为 50μl,同时加 2μl 无菌水到阴性对照管中,加 2μl 已知 DNA 到阳性对照管中。最后各管加 1 滴(约 30μl)石蜡油,快速离心数秒;如果使用有热盖的扩增仪,可以不加石蜡油;

(9) 将离心管放入热循环仪中。根据扩增基因座,设定 PCR 反应的热循环参数,进行扩增。

复合扩增较为常用的热循环参数:95℃ 60s,59℃ 60s,72℃ 60s,总共 28 个循环,然后 60℃ 延伸 45min,最后温度降至 25℃,直至取出反应管。如果反应体系中加的是金牌 Taq DNA 聚合酶,在进入循环前要求 95℃变性 11min。这个程序也可以扩增复合体系的单个 STR 基因座。基因座不同,PCR 反应条件也可能不同,因此根据所分析的基因座,具体选择反应条件。反应体系也可以减少到 10μl,各组分用量需相应减少。

12.3 PCR 扩增产物的分离与检测

12.3.1 DNA 片段的分离

1. 扩增产物的分离方法

由于 STR 基因座的重复单位短,等位基因片段间的长度差异很小,一般选用聚丙烯酰胺凝胶电泳分离,也有少数文献报道应用高浓度的琼脂糖凝胶电泳。

STR-PCR 扩增产物分离方法主要有:①水平不连续非变性聚丙烯酰胺凝胶电泳;②垂直非变性聚丙烯酰胺凝胶电泳;③垂直变性聚丙烯酰胺凝胶电泳;④高压液相色谱(HPLC);⑤毛细管电泳(CE);⑥基质辅助激光解析电离—飞行时间质谱(MALDI-TOF)分析。各个方法各有特点,表 12-16 为其分析速度和仪器花费比较。

表 12-16 STR 等位基因检测方法比较

技术	速度/样品	仪器花费
聚丙烯凝胶电泳法和银染法	数小时	+
自动测序仪	数小时	+ + +
毛细管电泳	数分钟	+ +
MALDI-TOF 质谱法	数秒	+ + + +

对于像 TH01,vWA,FES/FPS 等多态性稍低的基因座,一般可选用水平非变性凝胶银染系统,其优点是操作简便、费用低;对于像 ACTB 基因座具有许多等位基因内部序列变异体,要求较高实验技术以获得好的重复性,一般采用垂直变性凝胶体系结合荧光检测或毛细管电泳荧光检测系统;对于复合 STR 基因座体系宜采用垂直变性聚丙烯酰胺凝胶电泳,以增加分辨率。目前,一般使用含 6~7mol/L 尿素的变性聚丙烯酰胺测序胶(T4%~8%,C2%~5%,胶厚 0.2~0.5mm)。在这种凝胶中,DNA 片段的迁移率只与其长度有

关，而与其碱基组成无关，具有极高的分辨率，可以区分开相差1bp的二个DNA片段。

2.DNA片段检测方法

根据PCR反应中引物或底物是否带标记或修饰，STR-PCR扩增产物分离后的谱带显现方法分为三大类：

（1）银染法 用于常规的聚丙烯酰胺凝胶电泳分离扩增产物的谱带显示。PCR反应中引物或底物dNTPs不经任何修饰，PCR扩增产物不带任何标记。该法现被广泛使用，其灵敏度可达pg水平，所需试剂常见易得、廉价而且无害，整个染色过程只需1h左右。

（2）同位素、非同位素标记法 在PCR反应时将引物或底物dNTPs用放射性同位素标记，PCR扩增产物带有放射性，电泳分离后采用放射自显影的方法显现结果。这种方法灵敏度很高，结果显示在X光片上，本底干净，谱带清晰。其缺点是检验过程中存在放射性污染等问题。非同位素标记有辣根过氧化物酶、碱性磷酸酶和生物素等，电泳分离后进行酶联化学反应显现谱带。早期曾有文献报道此类方法，目前已基本不用。

（3）荧光法 在PCR反应时将引物或底物用特殊的荧光染料标记，使扩增产物片段带上荧光标记，电泳分离后用荧光扫描仪检测或PCR产物直接在全自动DNA测序仪上进行扩增片段的分离与检测，结果以扫描峰的形式显现出来，这种方法的灵敏度和准确性都很高，可以进行定量分析，但需要专门的仪器和试剂，成本较高。

对于单基因座STR分析也可采用MataPhor琼脂糖（FMC公司）电泳结合SYBR Green染色。

目前在法医DNA分析中STR扩增产物分离和检测方法上，应用的主要是垂直聚丙烯酰胺凝胶电泳银染法，荧光标记检测法和毛细管电泳荧光标记检测法。

12.3.2 垂直聚丙烯酰胺凝胶银染法

聚丙烯酰胺凝胶电泳银染法具有简便、快速、灵敏和成本低等优点，但只适用于单个基因座或几个扩增产物片段长度不重叠的基因座的复合扩增体系。

STR基因座等位基因片段长度差异比小卫星VNTR要小很多，为了区分开各等位基因，提高分辨率，一般需采用比较长距离的超薄聚丙烯酰胺凝胶。银染这种长而薄的凝胶，不宜直接将胶剥落在染色盘中染色，因为游离的大胶在染色过程中容易破裂，增加操作难度，必须将凝胶固定在一载体上，如将凝胶直接结合在玻板上，或结合在通过特殊处理过带有粘胶剂的聚酯薄膜上，再进行银染，这样凝胶不变形，操作方便。聚酯薄膜经过粘合剂Saline 174A处理，将其一面带有胶粘合剂，使其能与凝胶结合。凝胶结合在超薄透亮的薄膜上，操作方便，染色结果可以干燥后永久保存，避免了像固定在玻璃板上那么复杂繁琐，而且必须配备很多玻璃板的麻烦。

1.试剂：

粘胶剂（30μl）

硅化剂（10ml）

95%乙醇0.5%乙酸溶液：1ml冰醋酸加入199ml 95%乙醇

固定液（10%乙酸）：100ml冰乙酸加900ml超纯水

染色液：1g硝酸银，1.5ml甲醛，溶于1 000ml超纯水

显色液：30g无水碳酸钠，200μl硫代硫酸钠溶液（10%），1.5ml甲醛，溶于1 000ml

超纯水、10℃预冷，用前配制

尿素

丙烯酰胺母液（Acr:Bis = 19:1）

过硫酸铵

TEMED

10×TBE 缓冲液

2. 仪器

电源（如 LKB 或 Bio - Rad 公司）

长垂直聚丙烯酰胺凝胶电泳槽（20cm 以上长）及相应的玻璃板一套，如 Life Technique 公司的，或 Bio - Rad 公司或国产的玻璃板，大小各一块。

3. 操作程序

（1）聚丙烯酰胺凝胶制备前玻板的处理

① 短璃玻板的处理　取 2μl 粘胶剂溶于 1ml 95%乙醇 - 0.5%乙酸溶液，用镜头纸按水平方向和垂直方向涂擦玻璃板的一面，使粘胶剂均匀地涂在玻璃板每一处。晾干备用。

② 长玻板的处理　取 1ml 硅化剂用镜头纸均匀涂于玻板上，方法同上，5 ~ 10min 后用乙醇冲洗，晾干备用。

硅化的玻板使用两次后需再硅化；粘胶的玻板每次使用前均需用粘胶剂处理。

③ 将长玻璃板置于水平台上，硅化面朝上，放好夹条（spacer），将短玻璃板压在夹条上，含粘合剂面向内，用夹子夹紧两玻璃板或胶带封住玻璃板三边，调平，准备注胶。

另一种处理方法是将 PAG 固定在带粘合剂薄膜上，方法如下：

① 短玻板的处理　取 1ml 硅化剂用镜头纸均匀涂于玻板上，5 ~ 10min 后用乙醇冲洗，晾干备用。或用高压润滑油涂擦玻璃板，作用同硅化剂。

② 长玻板的处理　在长玻璃板中间用洗瓶滴上几滴蒸馏水，将聚酯薄膜放在玻板上，注意薄膜只有一面含粘合剂，粘合剂面朝上，薄膜的四角与玻璃板的四角对齐，并去除它们之间的气泡。

③ 然后放好夹条，将短玻璃板压在夹条上，硅化面向内，用夹子夹紧两玻璃板或胶带封住玻璃板三边，准备注胶。

（2）凝胶的制备

按表 12 - 17 配制 T4%、C5%、7mol/L 尿素的变性 PAG。

表 12 - 17　7mol/L 尿素、0.5×TBE 的 30ml 变性 PAG 组分

组分	4%	6%
尿素	12.6g	12.6g
超纯水	16ml	14.5ml
10×TBE	1.5ml	1.5ml
40%丙烯酰胺溶液（19:1）	3ml	4.5ml
总体积	30ml	30ml

待尿素完全溶解后，加入 10%过硫酸铵 120μl、TEMED 20μl，混匀，用平推法或注入法灌胶，插入加样梳子，最后用夹子夹好玻璃板，室温聚合 1h 以上。

平推法：将长玻璃板平放在水平台上，配以配套的夹条。一手拿住短玻璃板的一端，将短玻璃板的一端搭在长玻璃板的右端，另一手拿盛凝胶液的烧杯，缓缓将胶液倒在玻璃板一端，让胶液自动流到长玻璃板的另一末端，充满两玻璃板之间的空间，然后右手缓缓将短玻璃板向前推动，一边推，一边倒剩余的胶液，直至两玻璃板的左端（电泳时的阳极段）对齐，胶液充满整个空间，插好梳子。

注入法：用加样器或直接将胶液加在已组装好的玻璃板中，一边加，一边敲打玻璃板，使胶液流动加快，不易产生气泡，灌好胶后，室温聚合 1h 以上。

（3）STR 扩增产物的电泳分离

①预电泳及上样　待凝胶聚合后，将梳子拔出，用0.5×TBE 缓冲液冲洗加样槽（孔）以去除槽中的尿素，将鲨鱼齿梳子插入凝胶中 1～2mm，将凝胶安装在电泳槽中，上下电泳槽中加入 0.5×TBE。接通电流，在 40W 下预电泳至胶温度达 50℃左右（约 10min），同时取 4μl PCR 产物与 2×载样缓冲液等体积混和制成上样混合液，95℃变性 2min，立即放入冰水浴，上样前再冲洗加样孔以除去析出的尿素，取 5μl 上样混合液加在凝胶泳道中，在相邻泳道内加等位基因分型标准物。

② 电泳　在恒功率下电泳。对于长 32 厘米的凝胶，设定恒功率 40W 电泳约 80min、40 厘米的凝胶在 60W 恒功率电泳约 80min。具体时间可根据染料迁移位置确定。在 4%的变性胶中，溴酚蓝的移动位置相当于 40bp，二甲苯青蓝的移动位置相当于 170bp。在 6%的变性聚丙烯酰胺胶中，溴酚蓝的移动位置相当于 25bp，二甲苯青蓝的移动位置相当于 105bp。

（4）银染显色

①电泳结束后，小心将两块玻板剥离，将粘有胶的短玻板放在染色盘中；

②加固定液至没过凝胶，室温下轻轻摇动固定 20min。用去离子水洗涤 3 次，每次 2min；

③加染色液至没过凝胶，室温下摇动染色 30min 后，用去离子水冲洗 10s；

④凝胶于显色液中显色，条带清晰后，用 10%乙酸溶液停显；

⑤晾干，观察和记录结果。

带胶玻板于 10% NaOH 溶液中浸泡过夜，凝胶从玻璃板上脱落下来，用水冲洗玻璃板，干燥备用。

对于固定在聚酯薄膜上凝胶，银染过程原则上相同：

① 小心将短玻璃板揭去，将粘有胶的薄膜放在染色盘中；

② 凝胶在 20%乙醇固定液固定 5min，用去离子水漂洗 2 次；

③ 加 1.4%硝酸溶液到有凝胶的染色盘中，轻轻晃动 2min 后，用去离子水漂洗两次，每次 10s；

④ 加银染染色液（0.2%硝酸银溶液）染色 2min，用去离子水漂洗两次，每次 10s；

⑤ 凝胶在显色液（9g Na_2CO_3、170μl 甲醛/300ml H_2O）中显色，条带清晰后，用蒸馏水漂洗凝胶两次，每次 10s，然后用 10%乙酸溶液停显，干燥，得到可永久保存的凝胶图

谱（图 12－19）。

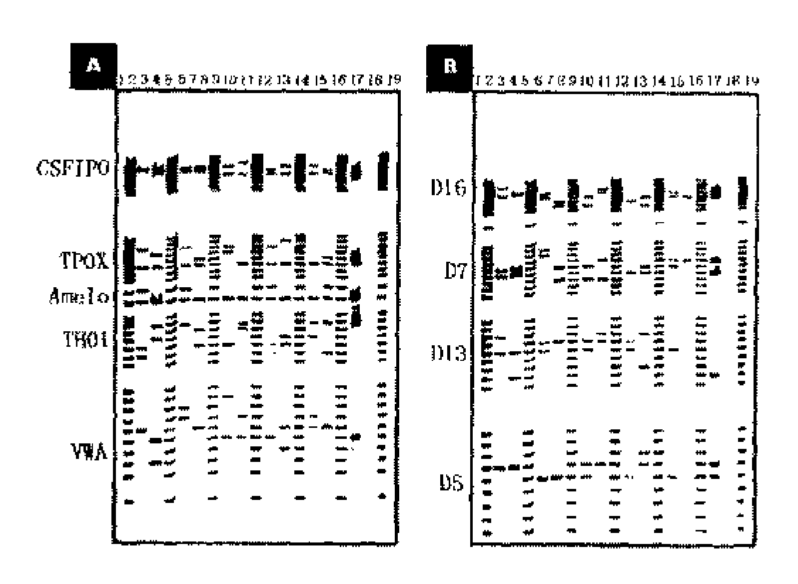


图 12－19 复合 STR 基因座扩增检测银染结果

4. 操作注意事项

（1）须用无核酸酶的去离子水稀释引物，否则引物会降解。

（2）玻璃板要清洗干净，以防灌胶时产生气泡，长玻璃板要硅化彻底，否则胶容易粘在板上。涂有粘胶剂的玻板要涂抹均匀，否则胶与玻板结合不牢固而剥落下来。按电泳槽说明书安装好灌胶架、玻璃板的底角保持平整。

（3）灌好胶后立即插入梳子，勿使梳齿下留有气泡。

（4）胶在显色液中时间太长或聚丙烯酰胺凝胶的质量不好会引起胶黄或背景太深。解决的方法是，加显色液前在倒掉硝酸银时将其甩干，放入无离子水中使胶板的正面和背面都洗干净，使用高质量的聚丙烯酰胺凝胶。

（5）聚丙烯酰胺凝胶的质量不好，或电泳温度太高，或样品量太多会使带模糊。解决的方法是，使用高质量的聚丙烯酰胺凝胶，胶溶液应贮存在棕色试剂瓶中。电泳温度控制在 40～60℃，减少样品量。

（6）STR 基因座重复序列较短，各等位基因片段长度差异小，相邻等位基因片段有时难以辨别，横向比对有时易产生偏差，再加上边缘效应或谱带漂移也会引起等位基因片段分型出现误差，因此要求每隔 3 个待检样品加一等位基因分型标准物，并且外侧的两个泳道也都加等位基因分型标准物。

（7）某些基因座的扩增产物变性后的二条单链由于碱基组成有差异，构象有所不同，在凝胶电泳中电泳迁移率不同，在凝胶图谱上表现为一个等位基因有 2 条带，会干扰结果判读。

（8）硝酸银溶液处理凝胶后，洗胶时间超过 20s，结合在 DNA 上的银被洗掉了，染色太浅，使 DNA 谱带模糊或无条带。

（9）染色液或显色液分解，谱带染成黄棕色，背景呈黑灰色。

（10）显色液温度过高，显色时间过长，造成染色背景太高。

（11）在银染步骤摇动不均匀，或在显色前的漂洗步骤不当，胶表面出现黑色的漩涡状图案。因此，在染色时必须轻轻摇动，在加显色液前用超纯水漂洗 20s。

（12）长胶板被粘合剂污染或硅化液处理不彻底，胶粘在两玻璃板之间。从短板上擦去多余的粘合剂，小心不要污染长板，放置时与长玻璃板分开；彻底硅化长玻璃板。

12.4 荧光标记 STR 基因座自动分析

12.4.1 概述

1. 荧光染料标记显带原理

荧光标记 STR 基因座自动分析是利用测序仪同时分离、检测带荧光染料标记的 DNA 片段。荧光染料由于含有共轭双键易吸收激发光，被激发到高能态（激发态），激发态极不稳定，当跃迁回到基态时，多余的能量以发光的形式释放出来（图 12－20）。用仪器如光谱仪、荧光扫描仪可以记录发出的光信号。将荧光染料标记在 DNA 片段，根据片段所带的特征荧光受激发后发出的荧光信号可以检测 DNA 片段。

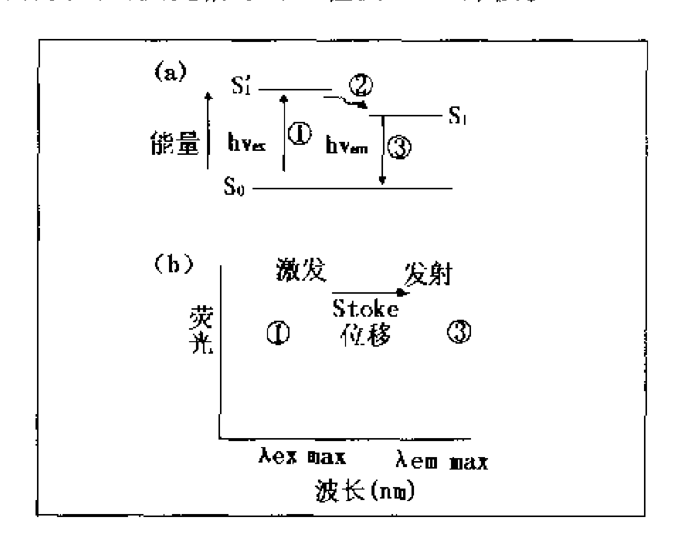


图 12－20　荧光发光原理示意图

(a) 激光产生的光子（hv_{ex}）激发基态（S_0）的荧光染料分子到激发转移态（S'_1），荧光分子发生构象变化，形成亚稳态（S1），回到基态，多余的能量以光的形式释发出来 (b) 荧光激发与吸收光谱

能量与波长彼此相关，吸收光谱与发射发光谱之间差异叫“Stoke 位移”（stoke shift）这种位移越大，越能用适当的滤镜将激发光与发射光分开。因此，染料的“Stoke 位移”越大越好。

荧光种类很多，用于 DNA 分析的荧光必须满足以下要求：（1）荧光素的吸收光谱应在激光束所提供的可见光区内（400～600nm）；并应最大限度地避免由凝胶板可能产生的干扰；（2）荧光基团发射光谱波长间应有足够的差异；（3）发射出的荧光应有足够的强度以达到检测器的灵敏度；（4）荧光基团的存在不应影响 PCR 延伸反应中诸如引物退火和 dNTP底物掺入等反应过程；（5）这些荧光基团，不改变 DNA 片段在电泳中的迁移（mo-

bility）速率，或者迁移速率的变化保持相对一致。

目前常用的荧光染料有 FAM、HEX、NED、ROX 、CY3、CY5 等，它们激发波长和吸收波长见表 12 - 18。图 12 - 21 为部分荧光染料分子结构。

表 12 - 18　各荧光染料的吸收波长、激发波长和颜色

染料	激发波长 Excitation max.（nm）	吸收波长 Emission max.（nm）	化学名称
6 - FAM	495	535	6 - carboxy - fluorescein
TET	519	539	4, 7, 2', 7' - tetrachloro - 6 - carboxy
HEX	529	560	4, 7, 2', 4', 5', 7' - hexachloro - 6 - carboxyfluorescien
NED	553	574	
ROX（CXR）	580	605	6 - Carboxy - X - rhodamine
Texas Red	578	602	
Rhodamine X	574	595	
TAMRA（TMR）	559	578	N, N, N', N' - Tetramethyl - 6 - carboxyrhodamine
CY3	552	565	
CY5	650	667	
JOE	526	548	6 - Carboxy - 2', 7' - dimethoxy - 4', 5' - dichlorofluorescein
Texas Red	587	620	
FITC Fluorescein	490	520	
TetraMethyl Rhodamine	546	572	
7 - Amino - 4 - Methyl - Coumain	354	445	

2. DNA 片段中荧光染料掺入方式

荧光染料标记 DNA 片段的方式有二种：（1）在引物合成时，将荧光染料连接到其中一个引物（一般选正链）的 5'端，使随后的 PCR 扩增产物中其中一条链的每一分子都带有一个荧光染料分子（图 12 - 22A）。在变性凝胶中，只显现一条带，结果清晰；（2）将荧光染料标记在 PCR 反应的底物 dNTP 上，在反应中，带荧光的 dNTP 在 DNA 链延伸时使每个片段标记上荧光染料（图 12 - 22B），这种方式使 PCR 产物的两条链都会带有荧光染料，在检测时会出现银染中出现的双带现象，干扰带的分辨与结果分型，在 STR 分析中不常应用。

3. 荧光标记 DNA 片段检测途径

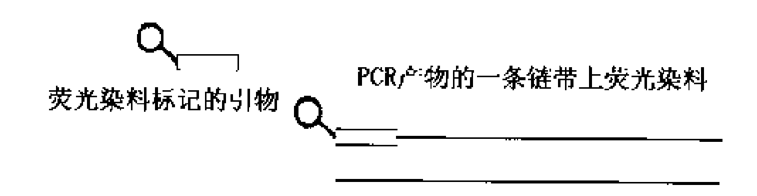
荧光标记 DNA 片段检测有两种途径：（1）电泳分离与荧光扫描分开的自动化程度较低的荧光扫描系统；（2）利用 DNA 自动测序仪检测，电泳分离与荧光扫描同时进行的荧光自动分析。

在凝胶电泳分离结束后，扫描检测带有荧光标记的 DNA 片段的系统叫荧光扫描系统。荧光扫描系统有双波长扫描仪，如 Hitatchi 的 FMBIO Ⅰ 型（图 12 - 23）和 Molecular Dynam-

图 12-21　部分荧光染料分子结构

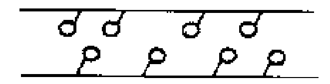
ics，可以分辨二个不同波长的荧光，它利用一种荧光染料标记标准分子量（内标），另一种荧光染料标记待测的 DNA 分子，还有多波段扫描仪，如 Hitatchi 的 FMBIOⅡ型能分辨四个波长的荧光，三种不同颜色道（蓝、绿、黄）可用于基因座片段，留下第四道（通常为最长波长的）作为内标。普通电泳后，扫描凝胶中的 DNA 片段，将 DNA 片段的大小与带强度根据荧光吸收峰的位置和大小被确定下来，转化成数字化的信息。荧光扫描仪将电泳和扫描独立分开，使电泳可以在普通电泳仪上进行，不占用荧光扫描仪，适用于大批量样品分析。其缺点是自动化程度低，分析中常采用同一时间荧光扫描，为了使所有等位基因片段均保留在凝胶中而未被迁移出凝胶，使一些基因座的等位基因可能因迁移距离短还未完全分离开，降低了分辨率与信息量。

DNA 自动测序仪（如 ABI310 系列、377XL）的激光荧光分析将 DNA 片段电泳分离与荧光扫描结合在一起，在电泳过程中即时记录 DNA 片段的荧光强度与迁移时间。这种分



A:引物掺入法

荧光dNTPs被掺入到PCR产物两条链内



B:底物掺入法

图 12-22 荧光染料掺入方式

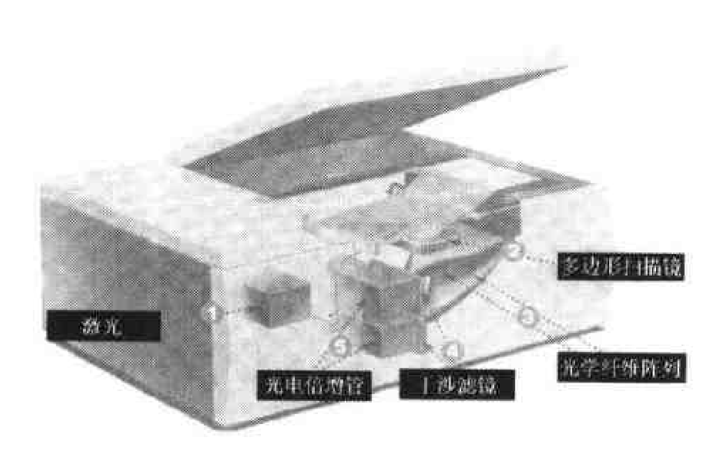


图 12-23 FMBIO I 型扫描仪

析系统自动化程度高。

4. 荧光标记自动分析

荧光标记自动分析 STR 基因座是利用荧光标记的引物在 PCR 扩增时，使 PCR 产物的一条链带上荧光标记。这种带有荧光分子的 DNA 片段在凝胶中从阴极向阳极迁移，按片段长度大小排列，当迁移到阳极端的激光扫描仪的扫描窗口，荧光染料受到激发，发出一定波长的荧光，按荧光强度记录下来，每一个带荧光染料的 DNA 片段电泳轨迹按各自通过激光扫描窗口的实际时间（real time）被记录下来，以荧光吸收峰来表示每一个片段(见前图 8-7)。峰值越高，表示该片段量越多；峰出现的时间与片段大小有直接关系，片段越小，峰越早出现。计算机保存所有片段通过激光扫描窗口的实际时间及其荧光特征，然后，根据同一泳道内标准分子量（简称为内标）的迁移率得到每一泳道迁移率的标准曲线，计算出待测样品的分子量大小，其精确度为 0.5bp。利用 Genotyper 软件或人工将测定样品片段大小与同一凝胶或一批加样的等位基因分型标准物比对，进行基因分型。

测序仪可以分辨四色荧光，实验样品可以用三种染料标记，剩下的一种荧光颜色（一

般为红色的 ROX）用来标记同一泳道内的标准分子量片段。最新的研究成果可以分辨五色荧光，四种荧光用来标记基因座，提高了复合扩增体系内基因座数目。

分析片段大小的第一台扫描仪为 362 基因扫描仪，这个仪器使用测序仪 ABI 370A 相似的结构，安装在潜水琼脂糖凝胶下面，用于分析片段大小为几百到数千碱基范围的 PCR - RFLP 产物和小卫星 VNTR 标记片段。随着微卫星 STR 标记的开发与应用，要求扫描仪的分辨率达 1bp，扫描仪 362 被改进为 373A 模式，及带 CCD 镜头的 XL377 基因扫描（图 12 - 24）。陆续开发出的以毛细管电泳为基础的低加样量的测序仪，如 ABI310，也采用 377 相同的 CCD 元件，进行自动加样，样品逐个分离检测、非常适合小型的、案件量较少的实验室使用。ABI3100 配有 16 道毛细管，MegaBACE 2000 和 ABI3700 分析仪配有 96 道毛细管，同时可以分析 96 个样品，提高仪器分析效率。

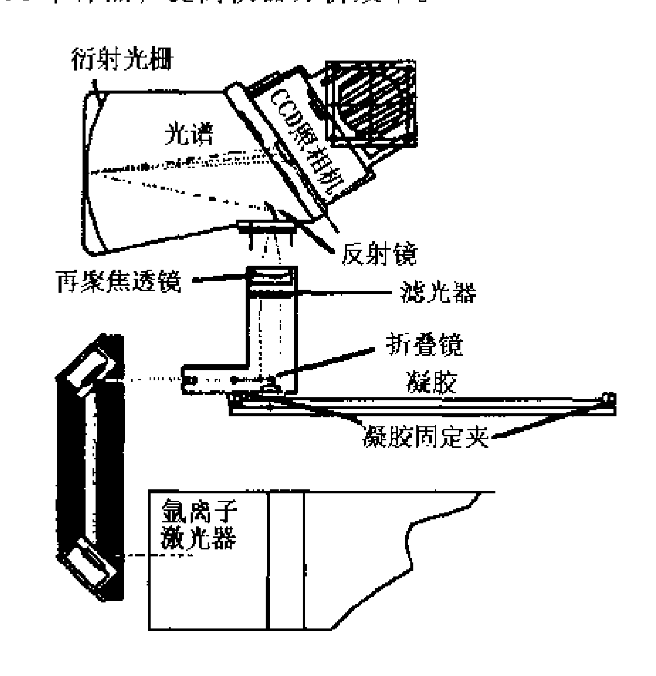


图 12 - 24　扫描激光的激发与光倍增仪检测系统的基本原理

5. 荧光标记自动分析系统的优点

与银染显色等非荧光标记自动分析系统相比，荧光标记自动分析系统具有如下优点：

(1) 使用内标，结果量化、准确、可靠，分子量测定可以精确到 0.5bp；

(2) 使用内标，消除泳道与泳道之间的相对误差；

(3) 能够分辨四色荧光，即使复合检测系统中有等位基因片段大小重叠的基因座也能同时得到检测，一次检测可获得较多基因座的遗传信息；

(4) 荧光激光自动扫描，不需要电泳后的谱带显现，高效率、省事；

(5) 荧光检测将结果转化成数字信息，适合随后的等位基因自动分型。

12.4.2 自动分析方法

12.4.2.1 垂直聚丙烯酰胺凝胶电泳荧光标记检测法

1. 试剂

(1) 水：对所有试剂、凝胶、缓冲液，使用 18MΩ的纯水或相当质量的去离子水。

（2）丙烯酰胺：LONG RANGE 公司的 50%丙烯酰胺溶液贮液，或 Bio-Rad 的 19:1 聚丙烯酰胺溶液，室温保存。

（3）尿素：分析级

（4）Tris/硼酸/EDTA 缓冲液：使用分析纯级。EDTA 必须是二钠盐，pH 为 8.3±0.1。

（5）TEMED：分析纯，避光保存。

（6）过硫酸铵：分析级，配置 10%（w/v）的贮液，要求在数周内用完。

（7）甲酰胺：去离子，冰冻保存。

（8）PCR 引物：可以自己合成或购买，每一个基因座只有一个荧光染料标记的引物。引物最好是 HPLC 纯化，分装后保存在 -20℃，防止反复冻融引起的降解。三种荧光素染料在合成过程中标记在引物的 5′端上。

（9）分子量内标：ROX™500

2. 方法

前面章节已详细介绍了 DNA 提取和 PCR 扩增等，本部分主要介绍仪器的准备，样品上样，电泳及下载原始数据的分析。分析过程分为四个主要部分：（1）仪器和样品准备；（2）样品上样与仪器运行；（3）数据分析；（4）基因自动分型。在此介绍的方法与程序，主要是针对 ABD377XL 测序仪而言。

（1）仪器和样品准备

① 凝胶的准备

自动测序仪上的测序或片段长度多态性分析的效果在很大程度上取决于电泳用的凝胶质量和玻璃板的洁净程度。按照仪器生产厂家推荐的有关试剂、玻璃板的清洁、凝胶液制备与灌注等技术的要求进行操作是很重要的。

玻璃板是自动测序仪的配件，质地均匀，对各种波长的光线通透性一致，先用滤纸沾去凝胶，用水浸泡玻璃板，用大量流水冲洗，直至玻璃板上不留任何杂质，最后用超纯水冲洗玻璃板，置于清洁通风处自然晾干。不要用纸巾擦干，以免在玻璃板表面留下擦痕和异物。

玻璃板干了以后，小心安装玻璃板，保证制胶模具包括夹条、梳子都正确配套。注意玻璃板的凝胶面朝里，玻璃板上带字的一面朝外，以保证检测的重复性。

ABI 373A 或 377XL 型使用的是变性聚丙烯酰胺凝胶，以 7mol/L 尿素作为变性剂。按所需凝胶浓度和用量配置凝胶，尿素必须完全溶解，用 0.2μm 的滤膜过滤并抽气处理，灌胶后插好梳子，在室温下放置 2h，使凝胶充分聚合。凝胶的制作参见前面章节。

② 仪器的准备和软件的设置

先开启测序仪和计算机电源，启动合适的基因扫描收集和分析程序，设置电泳参数，如收集方式、梳子大小、电泳时间等。电泳条件一般选为 3 000V 2.5h，或 1 000V 4h，根据所用的荧光染料，选择收集方式，确定滤光片种类，PE 公司的 PROFILER PLUS 及 COFILER 试剂盒使用 PLATE CHECK F 方式，Promega 公司的 1.2 及 1.1 试剂盒采用 PLATE CHECK C 模式。

关闭其他应用程序，保证计算机有足够大的内存空间。因为磁盘可用空间越多，系统越安全。如果 Macintosh 在网站上，应脱离开来，否则易发生“死机”现象，丢失数据。

③凝胶的安置与预电泳

凝胶的安置：用略湿的干净纸巾或滤纸擦去凝胶玻璃板外表面的凝胶、尿素结晶等，尤其是激光扫描口周围的脏物，否则凝胶的扫描基线会很高，干扰数据记录。清洁完玻璃板后，将凝胶玻璃板安装到机器上。点击“PLATE CHECK”进行玻璃板扫描检查，正常条件下凝胶及玻璃板的扫描基线低又平。如果基线不够平整，应取出凝胶组件，重新清洁表面。当基线平整达到要求时，将 1×TBE 电极缓冲液加在电泳槽中，用注射器吸取 1×TBE 液轻轻冲洗凝胶的加样孔，除去凝胶加样孔上面的尿素。预电泳 10～15min 使温度升至 50℃，同时使凝胶中的某些离子电泳出凝胶。

④样品准备

按甲酰胺：葡聚糖蓝（5mg/ml）：红色荧光内标 = 5：1：0.66 比例混合成上样缓冲液，取 1μl 上样缓冲液与 1μl 扩增样品混合成上样混合物后，95℃变性 2min 后移至冰水中骤冷，取 1μl 用于上样。葡聚蓝为示踪染料，以便加样时能看得清楚。甲酰胺为变性剂，结合加热使双链 DNA 变性成单链，随后迅速于冰水中防止复性，甲酰胺比重为 1.13g/ml，有利于上样。为了保证完全变性和混合液合适的比重，建议甲酰胺的终浓度为 50%或 50%以上。373 测序仪最多只能加样 4μl，377 测序仪最多加 1.5μl。如果不止一个 PCR 反应液加在一起，样品体积可能过大，可用真空离心或标准乙醇沉淀减少体积。

打开“SAMPLE SHEET”文件，按设计好的加样顺序编排样品表。

（2）上样、运行机器与收集数据

①样品上样停止预电泳，用缓冲液轻轻冲洗加样孔表面，除去尿素等影响加样的物质。将鲨鱼齿梳子插入凝胶中 1～2mm，保证梳子齿尖在同一水平线。制胶时将梳子平齐端插入胶中，待凝胶凝固后拔出梳子，用纯水冲洗凝胶，然后将鲨鱼齿尖端部分插入凝胶中。

按编排好的加样顺序将上述已变性的 1μl 上样混合液加入凝胶中。样品加样顺序要求案件相关样品不能加样在邻近泳道，以避免泳道与泳道之间的污染，如有可能，犯罪现场样品与嫌疑人样品在不同的凝胶上电泳分离，以防止因泳道间相互渗漏导致假匹配现象。

②机器运行再次检查设定的电泳参数，调入正确的样品表，选入正确的 matrix 文件（也可以不选 matrix）。点击窗口的“RUN”程序，命名并保存此 RUN 文件，机器开始电泳。

③数据收集

对于 373、Macintosh 计算机上的电泳时间设定必须与扫描时间设定一致或稍微超过一些，以保证在数据收集前没有数据输出。在电泳开始后几秒钟开始进行收集。

377 和 310 机器通过 Power PC Macintosh 操作，启动电泳运行数据存储自动进行。310 测序仪可在一个样品电泳时进行前一个样品的结果分析，377 则在电泳结束后一次分析所有样品数据。

在电泳过程中，凝胶基线和片段经过检测器的过程可以由扫描和图象窗口得到监测，在凝胶窗口观察从开始到目前的整个电泳情况，在电泳结束可以得到含全部等位基因片段的电泳图（图 12－25）。在早期 Macintosh 模式，如 373 型测序仪，必须在观察后关闭。

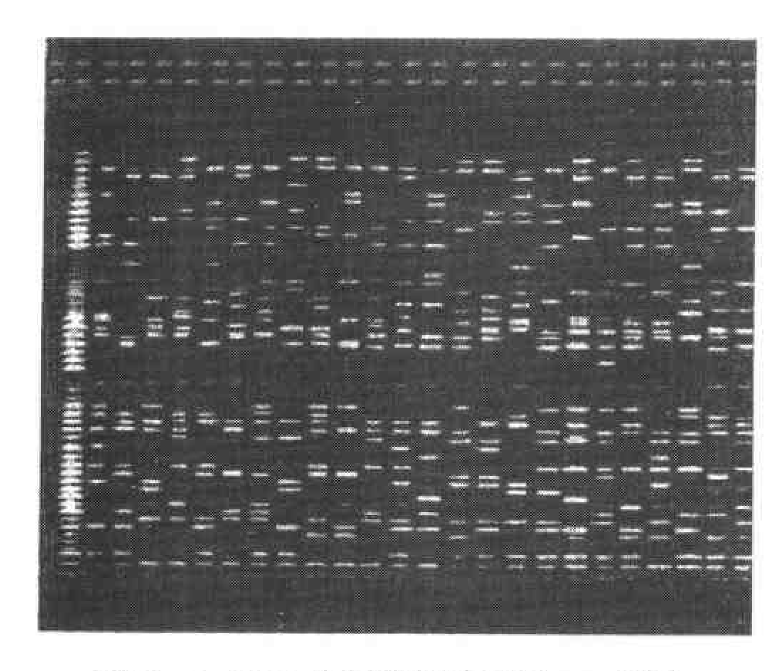


图 12－25　377XL 测序仪检测复合扩增 STR 电泳图

（3）数据分析

①校正追踪线　电泳结束后，打开“GEL”文件，调整电泳泳道追踪线，标定样品电泳踪迹。373 和 377 平板凝胶电泳的每一泳道电泳状态不同，可能出现带的弯曲或漂移，计算机自动划线与数据提取可能有偏差，需要校正。

②提取凝胶　划完线后，在“GEL”窗口选择“EXTRACT LANE”提取凝胶文件与样品表，提取后计算机自动显示分析屏幕。

③安装正确的“MATRIX”文件　在分析屏幕，点亮“SAMPLE FILE”栏，安装正确的“MATRIX”文件。在每次分析前，输入一个“MATRIX”文件，可以去除正常范围内的不同荧光之间的干扰。当一色道的信号很强以致饱和软件中的正常基线补偿（复合成分基质，matrix），在附近色道出现一个“pull－up”峰（关于“pull－up”峰将在后面介绍）。

④输入或定义内标分子量　在“SIZE STANDARD”选择正确标准分子量内标。ROX GS500 分子量大小为 75、100、139、150、160、200、250、300、340、350、400、490、和 500。

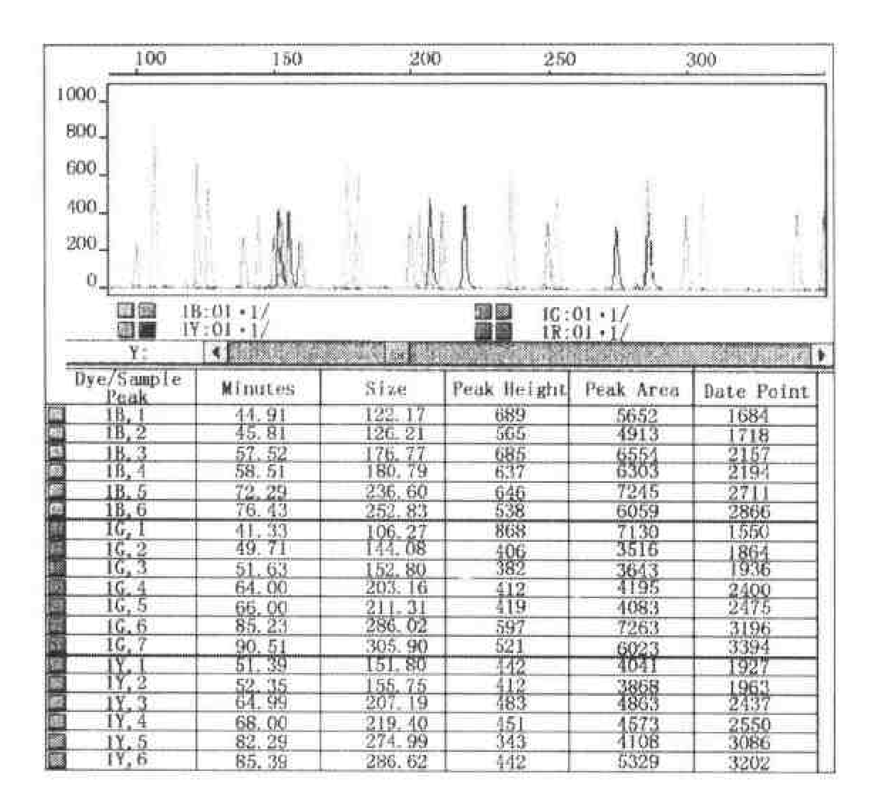
⑤自动分析　最后点亮蓝色、绿色、黄色及红色道，点击“ANALYSE”，机器开始进行自动分析，获得图 12－26 所示结果。

（4）等位基因分型

①基因扫描分析软件分析给出的每一个样品的结果是荧光谱带相对迁移率的电泳踪迹图及标明片段大小的表。根据所得到的等位基因片段分子量，与等位基因分型标准物 Ladder 大小比较，进行等位基因分型。

②内标检查　打开 GENOTYPE 软件，在“FILE”菜单中点击“IMPORT GENESCAN”，将基因扫描分析的结果输入，点击“CHECK GS500”检查内标，如果发现样品的内标不正确，需重新回到分析文件即带“π”文件，重新定义该样品的内标，重新分析。

③自动分型　双击左下窗口的“KAZAM”，机器开始进行自动等位基因分型标记。自



Dye/Sample Peak	Minutes	Size	Peak Height	Peak Area	Date Point
1B, 1	44.91	122.17	689	5652	1684
1B, 2	45.81	126.21	565	4913	1718
1B, 3	57.52	176.77	685	6554	2157
1B, 4	58.51	180.79	637	6303	2194
1B, 5	72.29	236.60	646	7245	2711
1B, 6	76.43	252.83	538	6059	2866
1G, 1	41.33	106.27	868	7130	1550
1G, 2	49.71	144.08	406	3516	1864
1G, 3	51.63	152.80	382	3643	1936
1G, 4	64.00	203.16	412	4195	2400
1G, 5	66.00	211.31	419	4083	2475
1G, 6	85.23	286.02	597	7263	3196
1G, 7	90.51	305.90	521	6023	3394
1Y, 1	51.39	151.80	442	4041	1927
1Y, 2	52.35	155.75	412	3868	1963
1Y, 3	64.99	207.19	483	4863	2437
1Y, 4	68.00	219.40	451	4573	2550
1Y, 5	82.29	274.99	343	4108	3086
1Y, 6	85.39	286.62	442	5329	3202

图 12－26　基因扫描分析软件分析的结果

动分型软件设定的自动正确分型的分子量误差为±0.5bp，当样品等位基因片段大小与LADDER相应的等位基因大小误差在±0.5bp范围内，计算机自动正确标定样品等位基因型，而对于误差超过±0.5bp，计算机就不能确定等位基因分型，显示“OL ALLELE ?”，此时需分析人员的介入，进行人工干预，比较样品等位基因片段与LADDER究竟相差多少，确定基因型。图12－27为自动分型结果。

（5）Matrix file 制作

Matrix file，是指用来修正荧光发射光谱之间迭加的一个数学基数（Matrix）。所用的四色荧光染料在同一激发波长下其发射波长相互之间有些干扰，在附近色道出现一个干扰峰（pull up peak）。Matrix file 在分析中校正光谱间的迭加，抵消不同荧光染料相互之间的荧光吸收干扰。Matrix file 对仪器及电泳条件是特定的。因此，当首次使用仪器、不同染料的或使用不同凝胶或不同的电泳条件时，必须用标准染料测定他们在不同吸收波长下的吸收，建立一个用于光谱补偿的 matrix 文件。Matrix file 的制作分9个步骤，图11－28为其流程图。

下面具体介绍 matrix file 的整个制作程序。

①matrix file 标准样的准备

对于ABD 310按每种染料编好样品管，取1.0μl matrix标准样与12.0μl去离子甲酰胺混合，轻轻振荡3～5s，快速离心混合液，保存于2～6℃。

对于ABD377　50mg/ml葡聚糖液与去离子甲酰胺按1:5混合成上样缓冲液，然后取2.5μl每种染料标准物与1.5μl上样缓冲液混合，制成上样混合液，在上样前每个样品在

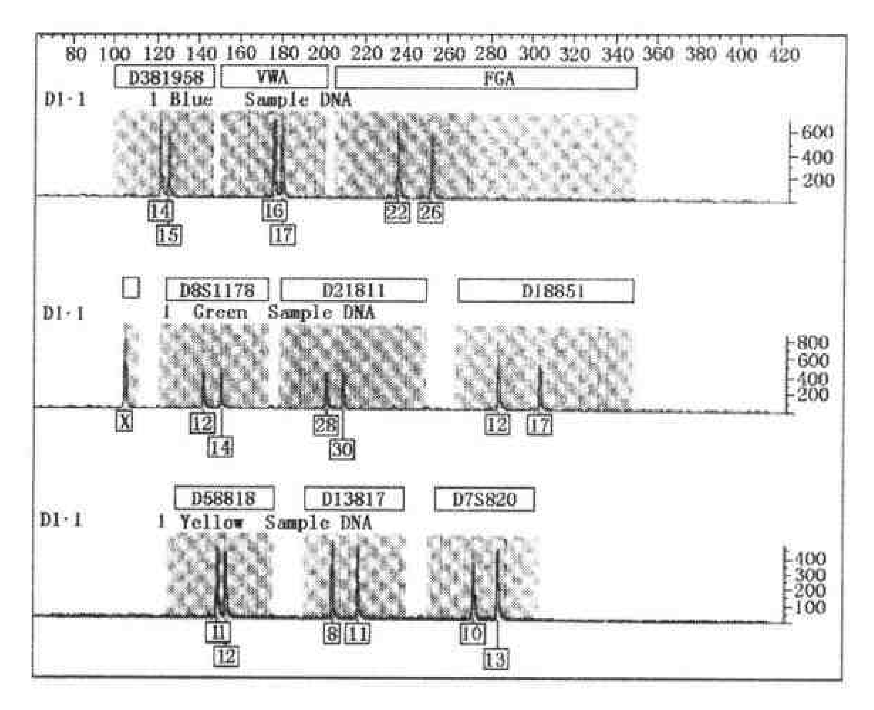


图 12－27 自动分型结果（三种不同颜色合在一起）

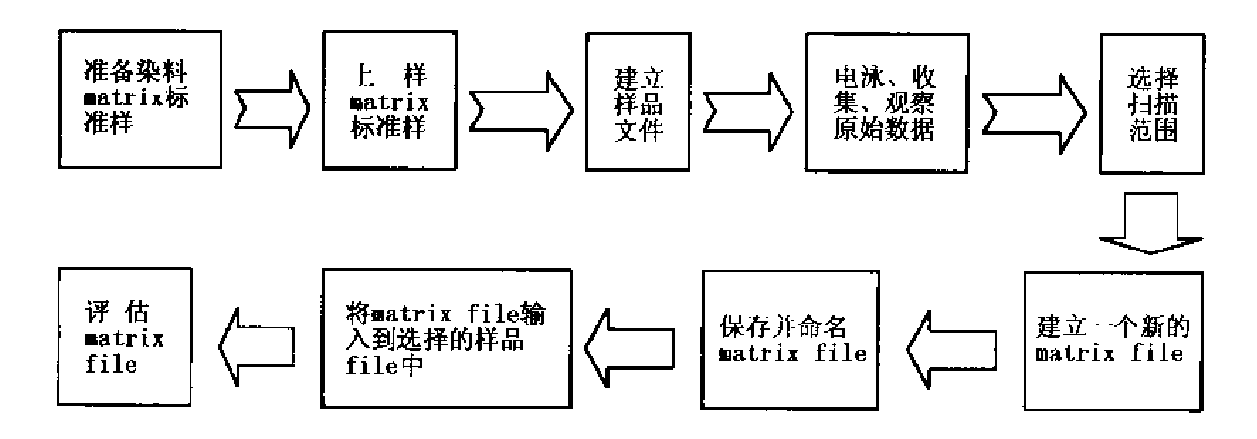


图 12－28 Matrix 文件制作流程图

95℃加热 5min。

②上样 matrix 标准样

③样品表与电泳、数据收集

对于 ABD 310：编好样品表与加样表，开始电泳收集数据。

对于 ABD377：编好样品表，将 0.5～2.0μl matrix 标准物隔道加在变性的聚丙烯酰胺凝胶中，选择合适的收集模式、电泳电压及时间等参数，确保选用正确预电泳和电泳模式，电泳，收据数据。

④检查原始数据 电泳结束后，在 file 菜单中选择“new”，在出现的对话窗中点击“project”，找到并打开所电泳的 matrix “run folder”，选择代表蓝、绿、黄和红染料的 4 个

样品文件，然后点击“add”和“done”，完成样品文件的转移。在分析窗口，选择4个matrix标准样的样品文件，在project menu中选择“Raw data”，此时出现电泳轨迹的原始数据。开始检查原始数据，要求四种matrix标准物均有数据峰、数据没有反常现象、基线平整稳定以及峰高在200~4000RFU范围之间。

⑤选择扫描范围　移动鼠标，去掉引物峰，根据起点和结束点的扫描线，确定扫描的起点和终点，记录扫描数目，关闭原始数据窗口。

⑥建立新matrix file　在“file menu”中选择“new”，出现一个对话窗口，在“number of dye”滚动窗中先确定所用的染料数目4（如用5色染料，则选为5），在各种染料相应的栏目中填入各自扫描起点的扫描数，根据输入数据点的总数来计算matrix，一般建立一个matrix至少需5个峰。然后点击“OK”，建立一个新的matrix文件。图12－29是计算迭加校正的数据，对于任一染料，在最适滤光器下的染料荧光值为1，其它情况下的值小于1，但等于或大于0.0010，这个数值即是系统必须补偿的迭加量（表12－19）。

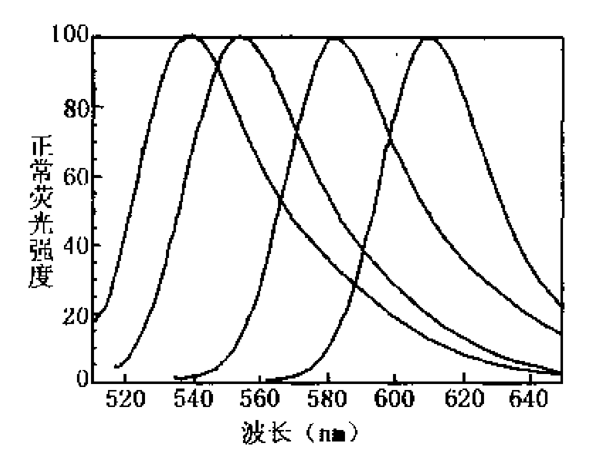


图12－29　不同荧光染料的吸收与校正数据

表12－19　荧光在不同波长下的吸收

	反应（310F的Matrix）			
	蓝色	绿色	黄色	红色
蓝色	1.0000	0.5423	0.0428	0.0010
绿色	0.6601	1.0000	0.5460	0.0059
黄色	0.3459	0.5735	1.0000	0.0893
红色	0.1693	0.3236	0.5922	1.0000

⑦matrix file的保存和命名　在“file menu”中选择“save”，将建立matrix file命名并保存在数据收集软件中，拷贝一份到“GS matrix folder”中。

⑧将matrix file调入样品文件中　待新的matrix file制作好后，将matrix file调入要分析的样品文件中。

⑨评估matrix file　建立了matrix file及调入到样品文件中后，下一步是对matrix file的评估。

分析用于做 matrix file 的样品文件：将 4 个染料的 matrix file 样品文件显示在同一屏幕中，每一个样品文件必须只有代表染料 matrix file 标准样的单一颜色的尖锐的峰出现在泳道中，所有其他颜色的线相对平滑而没有峰，表明 matrix 正好补偿了光谱的迭加，这样的 matrix 才可以应用。

12.4.2.2 ABI PRISM 310 毛细管自动分析

自 1995 年 PE Applied Biosystems 开发出 ABI Prism310 毛细管电泳（capillary electrophoresis）的基因分析仪以来，广泛地应用于 STR 基因分型，国内已有数十家刑事技术和司法鉴定机构拥有此仪器。

ABI310 是单道毛细管多色荧光检测的仪器。工作人员只需将毛细管安置在仪器中、把注满多聚体溶液的注射器安置在仪器上，将要分析的一批样品放到样品自动进样槽上，然后点击“RUN”就可以开始运行，大约在 30min 内完成一个样品数据的收集与分析。最大的优点是快速、自动化程度高、可以随时停机、随时调整分析样品的顺序。

1. ABI310 的结构

ABI310 基因分析仪的基本结构如图 12－30 所示。在进样口电极和泵活塞之间有一根毛细管，通过泵、活塞作用使毛细管填满多聚体溶液。加热板用来调节温度，使毛细管的温度恒定在一定温度。样品放在样品自动加样盘，在进样过程中上下移动将样品进入到毛

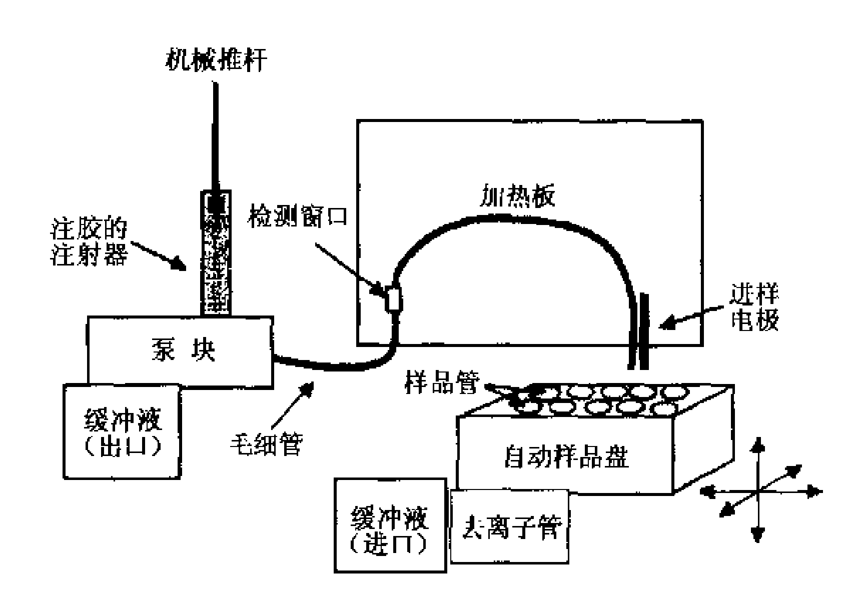


图 12－30 毛细管电泳结构

细管。毛细管的每个组成对毛细管电泳都十分重要，下面将详细讨论：

（1）毛细管

用于 STR 基因分型的毛细管内径为 50μm，长为 47cm。外层涂有一层聚亚胺保护层，使毛细管具有一定的柔性而不易折断。涂层不透光妨碍毛细管内的检测，因此在接触激光激发部位的毛细管去掉涂层，暴露出毛细管。

影响毛细管长度的两个因素是峰的分辨率与分离时间。一般，毛细管越长，分辨率越好，分离时间越长。

ABI310毛细管电泳系统的毛细管一般内壁没有用硅烷基团包埋，而用分离DNA的POP-4多聚体溶液线性包埋毛细管内层，防止电渗流。

毛细管是毛细管电泳中的主件，需要细心维护。电泳结束后，将毛细管的二端插入水或缓冲液中进行毛细管的维护和保存。如果毛细管二端干了，那么尿素或缓冲液中一些盐分会形成结晶。由于毛细管管窄，非常小的结晶或颗粒都会引起毛细管的堵塞，电流较低或甚至没有电流。因此，当较长一段时间不用时，毛细管应保存在去离子水中。

（2）多聚体溶液

用于310仪器STR分型的多聚体为POP-4。POP是perfomance optimized polymer的缩写。代表4%浓度的非交联的组成的。含有4%的单一多聚体，[二甲基丙烯酰胺（N，N-dimethylacryamide，DMA）]，8mol/L尿素，5% 2-pyrrolidinone，100mmol/L Tris-（hydroxymethyl）methyl-3-aminopropane-sulfonic acid（TAPS，pH8.0）。

（3）电泳缓冲液

电极缓冲液提供离子在毛细管内传递电流，如果使用不正确，电流会波动，影响DNA分离。PE公司提供的含有EDTA的10×Genetic ananlysis缓冲液，适合于STR分析。用时只需用9倍水稀释，配成1×缓冲液。1×缓冲液为：100mmol/L TAPs，1mmol/L EDTA（pH8.0）。

2. STR基因分型的ABI310的操作

在分析每一个样品前，ABI310毛细管系统必须作好准备，准备过程有以下4步：

（1）把毛细管放到仪器中，固定到检测窗口；

（2）将多聚体溶液加到注射器，赶走所有气泡；

（3）缓冲液槽中放好缓冲液，保证两个电极之间有电流流通；

（4）在管子中放入去离子水，以保证毛细管末端不干燥。

整个操作过程如下：

（1）开机 首先打开310主机电源。仪器正面的绿灯亮为正常，如红灯亮则表示仪器有问题，需与机器维修工联络。接着按住计算机键盘右上角△键一段时间以启动计算机，打开310 Data collection软件。

（2）进入310 Data collection之后，对310主机进行自动进样器的X，Y和Z的复位以及挤压泵的复位。此步在更换毛细管或更换电极、或关机后、或每次启动310 Data collection时必须进行的，如果不进行上述操作可以越过这步。

①自动进样器X，Y座标复位 在Window→Manual control→Function中（按[⧗]）选择Autosample home X，Y Axis→按Excute键，机器自动进行X，Y的复位。

②自动进样器Z座标的复位 在Window→Manual control→Function中选择Autosample home Z Axis→按Excute键，机器自动进行Z的复位。此步必须在①完成后才能进行，否则会损坏电极。

③挤压泵的复位 Window→Manual control→Function中选择syringe home→按Excute键，机器自动进行复位。

（3）安装电极

首次使用装置或电极严重变形时，需要安装新电极。要求带无粉手套操作，以免污染

电极。打开仪器门，再打开加热板，拧下装电极的基座，将电极放进电极孔中，定位钩放进大孔中，把基座拧回去，电极安装完毕。检查电极长度，如果太长，剪去多余部分，如果太短，则更换电极。

（4）安装毛细管

① 从左边开始安装毛细管，先将毛细管穿过活塞，向左穿，一直到十字交叉口处。毛细管头不要超过十字交叉口，以稍偏右为好。

② 将毛细管检测口部分安装在主机激光扫描口处，盖上检测窗盖子。

③ 用胶带将毛细管固定在加热板上，并把毛细管的另一端扎入电极塞中间的毛细管孔中，调整毛细管与电极的相对位置，使毛细管比电极略低 0.5cm。最后用胶带把毛细管牢固地固定在加热板上（三点固定）。

④ 拧紧毛细管左端的活塞，用手拔一下毛细管，以拔不动毛细管为准。

⑤ 标定毛细管与自动进样器的相对位置，每次换毛细管必须做此步。

打开仪器右边门，在 instrument 栏中选 “Autosample Caliberation” →按 “Start”，然后根据计算机提示的要求进行。拿走样品盘，并且拿走 3 号位置的离心管。然后按 “Resume”，进行前面点的标定，根据两者的位置，按计算机屏幕的方向键，使毛细管的底部尽可能接近金属点，金属点处于毛细管和电极中间，并靠近毛细管。校正后，按 “Set”，自动进样盘会自动推出，按计算机屏幕提示，把样品盘和 3 号位的离心管放回；按 “Resume”，进行后面点的标定，方法和要求同前面点。标定后，按 “Done”。

（5）灌胶

①将干净并干燥的透明有机玻璃的样品块安放到主机上。

②将 POP 胶从冰箱取出，室温预热。用少量 POP 胶液湿润注射器，注射器吸胶时要缓慢，避免气泡的产生。倒置或轻弹注射器把气泡赶出针筒。将注射器放到样品块上，拧紧。

③先灌下面通道：在 Window→Manual control→Function 中选择 “Buffer Valve Close”，按 “Excute”，机器自动关闭缓冲液阀门。关好缓冲液阀门，松开废液出口，压灌注射器，直到胶液出现在 “O” 型圈位置。然后，将废液出口拧紧。

④再灌右边通路：先松开毛细管活塞，压灌注射器，直到胶液出现在 “<” 字型槽内。再把活塞拧紧。

⑤最后灌左边通道：在 Window→Manual control→Function 中选择 “Buffer Valve Open”，按 “Excute”，机器自动把缓冲液阀门打开后，轻轻压灌注射器，胶慢慢灌进左边通道，直至胶出现在阀门的 “V” 字型凹槽内。在 Window→Manual control→Function 中选择 “Buffer Valve Close”，按 “Excute”，关闭缓冲液阀。如果通道内有气泡，必须继续灌胶，直至气泡出去为止。

⑥向右扳注射器固定器，然后选择 “Syringe down”，一点一点调整固定器位置，使注射器推杆顶部正好在固定器凹环内。

⑦选 “New”，选择 “Sample injection”，调入 “Sample sheet”，选 “fill capillary”，取消自动分析。完成后机器自动灌注毛细管。

（6）准备缓冲液

取 10×TBE 缓冲液 1.4ml，加 12.6ml 超纯水，混合，配成 1×TBE 缓冲液。将其中的

4ml 加到 1 号缓冲液杯中，并作好标记，以便与 2 号的水杯区分开；放在 1 号位。取剩下的 10ml 1×TBE 缓冲液到阳极杯中，并安放在样品块中。在 2 号杯中放入 4ml 的超纯水，放于 2 号位。取 1.5ml 无盖的离心管，加 1ml 左右的纯水，放于 3 号位，用来收集废胶。管内的水用来防止电泳结束后因毛细管暴露空气中而干燥。

（7）设置分析参数

运行 Genescan 分析 3.1 软件，在设定菜单（setting menu）中选择分析参数，设置如下：

分析范围（Analysis Range）：2400～7000

数据处理（Data processing）：选择基线（Baseline）与多组分（multicomponent）。

峰检测（Peak detection）：B：150、Y：150、G：150、R：150

最小峰半宽（Min peak half width）：3 pts

大小范围（Size call Range）：75～400 bp

裂开峰校正（Split peak correction）：none

校正范围（Correction limit）：30 Data pts

设置完毕，退出 Genescan－分析软件，运行 collection。

这些在机器开始使用时设定，以后不需要再设定。

（8）设置电泳温度

在 GS STR pop4（1ml）F 运行下，在手动控制下设定温度为 60℃，进样时间（一般为 5s～10s），以控制进样量。

（9）样品准备

按去离子甲酰胺：GeneScan－500［ROX］分子量标准物＝24∶1 混合两者，离心，取混合液 25μl，再加入 1.5μl（扩增产物）或等位基因分型标准物（allelic ladder）1.5μl，或按甲酰胺：GeneScan－500［ROX］：PCR 产物＝12∶0.5∶1 混合混匀，离心，置 95℃变性 3min，置冰水浴至少 3min。然后将样品放在样品盘中，准备电泳分析。

（10）设置样品表（sample sheet）

根据自动样品槽的孔数打开合适的 sample sheet，按分析的样品顺序填入到样品表，在 “pres” 处都打上 “×”，复制样品名到 ”sample info” 与 “comments”，标明荧光染料（蓝、绿、黄和红），每个样品的分子量内标和它的相应的颜色都要注明。

（11）编排进样表（injection sheet）、电泳，收集数据

编好样品后，点击 “project”，在打开 “GeneScan”，再打开 ”injection list”，在 “sample sheet” 栏中调入样品表。设定电泳时间和进样时间。电泳时间一般为 24min，对于使用次数较多的缓冲液，要求延长电泳时间到 26min，否则有可能有的基因座在设定时间内还未电泳通过激光扫描检测窗口，收集不到数据，进样时间一般为 5s。在进样表中，每个样品选择好收集模式。选择原则同 377 测序仪，对于使用 PE 公司试剂盒，模式是 “GS STR POP4（1ml）F”，默认：15kV 进样 5s，15kV 电泳分离 24min。

毛细管的加热板升温到 60℃，热平衡毛细管，激光转到满功率（～10mW）等正常工作后，机器自动分析每个样品。自动分析样品的过程：

①毛细管注胶（capillary fill）　加力给注射器上，向毛细管注入多聚体胶液。对于

POP－4 和 47cm 毛细管，每次进样向下移动 5～10 步。如果注射器移动超过 10 步，说明泵活塞口发生渗漏，或毛细管被堵住。

②预电泳（pre－electrophoresis） 分离电压提高到 10 000V，电泳 5min。这步检查毛细管内是否有气泡，平衡分离系统。如果毛细管内有气泡，没有离子通过毛细管，提高电压电流仍为零。

③水冲洗毛细管（water wash capillary） 毛细管反复吸取去离子水，除去缓冲液中干扰进样的盐。

④进样（sample injection） 自动进样槽移动到 A1 位置，向上移动使毛细管接触到 DNA 样品，准备进样。加压给样品，几 nl 样品进入到毛细管一端。默认值为 15 000V 5s。

⑤水洗毛细管（water wash capillary） 毛细管吸水数次，除去粘在毛细管外面的任何污染溶液。

⑥吸水（water dip）——毛细管放在水里吸几次。

⑦电泳与检测 自动进样槽移到入口的缓冲液瓶中，分离电压通过毛细管、进样的 DNA 样品通过 POP－4 多聚体胶液开始分离。当带荧光的片段通过扫描口时发出的荧光被记录收集数据。这时收集的为没有任何光谱处理的原始数据。

大约在 30min 内完成每个样品分析的整个过程，默认的电泳时间为 24min，毛细管注胶和预电泳步骤大约 6min，400bp 左右的 DNA 片段在 47cm 毛细管 15 000V 电压下电泳 24min 通过毛细管。

（11）分析数据

运行 Genescan 分析软件，标定内标 Genescan－500［ROX］，进行分析。点击电泳的某一样品文件以启动 Genescan 软件。在"file"内点击"new"，点击"project"中的"add sample files"，"add"或"add all"输入所要分析的样品，点击"done"，计算机自动出现分析窗口。然后，与 ABD377 分析一样，输入正确的 Matrix 文件，定义分子量内标，最后进行分析。需要注意的是，不要定义分子量内标的 250bp 片段，要跳过此片段。因为此片段在 310 仪器上实际测定不是 250bp。

（12）关机

①取下毛细管，使其两端分别置于两个装蒸馏水的 1.5ml eppondorf 管中；

②用去离子水清洁注射器；

③运行样品盘 X－Y 轴，Z－轴回归位；

④关闭电脑（先退出程序，后关主机）；

⑤关闭 310 主机开关。

与 ABD377 测序仪一样，在开始使用或改变条件时需制作 matrix 文件，方法见前。

310 毛细管电泳分析样品是顺序分析样品，而不是像平板凝胶电泳平行分析，目前已开发了多道的毛细管电泳，大大加速 DNA 分析速度。

3. 影响 310 分析 STR 基因座的因素

（1）样品进样

对于 ABI310 基因分析仪，DNA 样品通过电进样自动进入到毛细管内。电极和毛细管必须深深插入到样品管中，保证进样电压加上后有电流，STR 分析默认的标准进样为

15kV 5s。增加电压或延长时间可以增加进样量（例如从 5s 加到 10s 或 15s），或者透析除去溶液中的盐离子或者将样品溶于去离子水中可以增加信号强度。

电动进样量很大程度上依赖于样品中盐浓度。在一定时间内一定电动进样电压下 DNA 进入到毛细管的量主要受样品的离子强度影响。DNA 样品进入毛细管中的速度与样品中的离子强度成反比。因为与 DNA 分子竞争的盐离子越多，使进入毛细管 DNA 分子越少。图 12－31 为不同电导率甲酰胺稀释的样品分析结果，从图中看出，当稀释用的甲酰胺电导率为 1180μS 时，信号明显降低。用去离子水稀释 PCR 扩增产物或对样品进行透析，可以改变样品/缓冲液离子强度比，降低离子强度。

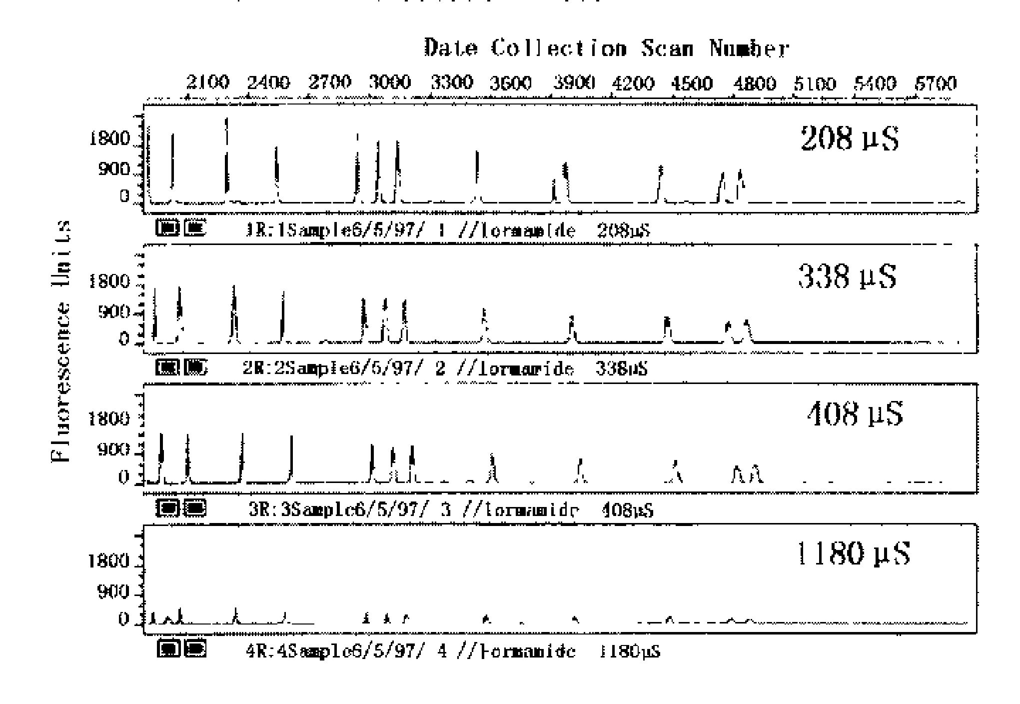


图 12－31　同一样品用不同甲酰胺稀释的结果（ABI310）

对于 ABI310，最佳的结果是样品的信号强度应在 150－6000 相对荧光单位（relative fluorescence units，RFUs），如果峰高超出范围（大约在 7500 RFUs），减少样品进样时间重新进样，比如 2s，可使峰高落在范围内。

（2）温度控制

室温的波动不稳，会影响 ABI310 测序仪分析准确性。在 60℃下进行 STR 电泳分离可减少 DNA 二级结构形成和链内杂交结构产生。当在较低温度下，DNA 的链没有完全变性而形成 DNA 二级结构，使 DNA 大小测定精确度下降；电泳温度变化将会导致分子量内标和 STR 等位基因片段的迁移率不同，使 STR 等位基因大小计算发生变化。

ABI 310 有一个加热板，用来升温和维持毛细管内的温度，但是不管怎样，毛细管的进口段和出口段总有几厘米暴露于空气中，没有直接与加热板接触，因此要求室温的波动幅度在 ±1℃，保证 ABI310 的 DNA 分离的准确性。

（3）缓冲能力

缓冲液使用一段时间后，缓冲能力会下降。在电泳过程中离子通过毛细管，阳离子聚

集在阴极端，阴离子聚集在阳极端。离子出现不平衡，使缓冲液缓冲能力下降。此时需要重新添加缓冲液或更换缓冲液。

（4）毛细管

分析一定数量的样品后，需要换用新的毛细管。厂家推荐的毛细管寿命为 100 次，实际一般为 500 次进样。因为毛细管内层的线性包埋不行了，噪音很高，不能分开二个相近的等位基因，出现不正常宽的峰。

毛细管从 ABI310 仪上卸下来，随后用水、四氢氟喃（tetrahydrofuran）、盐酸和多聚体溶液洗可以再生毛细管。也可以用泵泵进几 ml 去离子水，然后用 Tris - EDTA 缓冲液洗毛细管，除去结合在内壁上的一切物质。

12.4.3 操作注意事项

1. 实验中所有试剂包括水的改变都可能影响结果，必须注明所有相关资料，例如：厂家、批号、生产日期及购进实验室的日期等。一旦实验中发生任何差异，便可以根据这些资料很快检查、发现原因。

2. 引物是由荧光标记的，引物及 PCR 扩增产物必须避光保存。

3. 调节样品上样量，保证带的荧光强度落在检测器线性范围内（40 ~ 4000 RFU）。如果上样量太多，会出现 pull - up 峰，峰（带）较宽。对于较浓的 PCR 产物，可用超纯水 10 倍稀释产物。

4. 对于使用 ABD 377XL 等测序仪平板凝胶电泳检测的，玻璃板一定要洗干净，否则所灌制的凝胶不均匀，带易扩散或拖尾，分离效果差，而且凝胶易发胀，使谱带弯曲。如果在激光检测器处玻璃板没有洗干净，本底很高，会影响样品检测，结果不可靠。

5. 凝胶溶液要求过滤并脱气，凝胶中有不溶物或气泡影响凝胶的均一性，降低分辨率。

6. 样品加样前，一定要冲洗凝胶加样孔以去掉尿素和没有聚合的胶液，尿素剩留在凝胶孔处会使谱带变形扩散。

7. 样品必须完全变性（95℃，2min），变性后如果不马上加样，立即将变性后的样品放在冰水浴中以防复性。否则，内标出现杂带，样品出现额外峰，影响结果判定。

8. 平板凝胶电泳时要防止泳道间样品的渗漏，建议隔道加样；案件中相互比对样品不要加在相邻的泳道内，以防因渗漏而影响结果，最好是同一案件中的相关样品在不同凝胶上电泳。

9. 在每一凝胶泳道内加入分子量大小标准可以使等位基因片段大小准确到 0.5bp。但是，如果分子量内标不正确，那么随之测定的等位基因片段大小就会发生错误，等位基因分型不正确。因此在进行 Genotyper 软件自动分型前，必须检查内标，如果内标不正确，必须进行重新定义，重新确定片段大小。

10. 等位基因自动分型时，等位基因分型标准物 Ladder 的第一等位基因的分型有时会有误，导致分析样品等位基因的分型错误，必须检查一下 Ladder 分型是否正确。

11. 对于用 ABD 310 型毛细管测序仪，在使用前毛细管必须归位，用多聚胶液灌注毛细管时一定要将气泡排干净。

12. 在使用 POP4 胶液时，要预先将胶液从冰箱中取出，室温放置一段时间后，使之

温度回升到室温后才能灌胶，因为胶中有尿素，易析出。

13. 电泳过程中，缓冲液和纯水因蒸发而减少，需要不断加液，使液面在刻度线处。否则，液面太低，毛细管无法冲洗，电极接触不到液面，无法电泳。

14. 分析时使用另一凝胶中的等位基因分型标准物 ladder，等位基因自动分型中可能会出现许多“OL ALLELE”。这是因为两次的电泳条件不同，等位基因分型标准物 ladder 与样品的同一等位基因的分子量大小差异在 0.5～1bp，计算机不能确定其基因型，需要人工纠正。

15. 建立一个分析条件（仪器、试剂等）专一的 matrix 文件。进行远程计算机分析数据时必须注意使用与结果一致的 matrix 文件。

16. 对于荧光强度低于 150RFU 的峰，结果解释要慎重。

17. 样品变性时间不宜过短或过长，过长的变性使 DNA 片段拖尾，荧光峰变得钝而宽，降低分辨率，无法精确确定等位基因片段大小。

18. Genescan 分析前应仔细检查每个吸收峰，避免将 Pull－up 峰误认为吸收峰。

12.4.4 基因分型注意问题

不论什么案件，不论利用什么方法、技术检验分析物证，其共同的目的或最关心的是案件检验后的结论，结论是分析人员对检验结果作出的正确解释或分析，因此结果的解释十分重要。不论是银染系统还是荧光标记检测系统，如果对分析系统了解不深入，同样的实验结果，所作的结论可能有些差异。

1. 等位基因以外的额外带

单一一个 STR 基因座的 PCR 扩增结果一般是：纯合子个体只有一条带或一个峰，杂合子个体拥有两条大小不同的带或峰。出现的第三条带或峰被称为额外带或峰。荧光标记自动分析结果以一个个荧光峰表示等位基因片段，这里描述的峰即通常说的带，额外带或峰有以下几类：

（1）阴影带（“Stutter”带）或峰

阴影带（stutter band）或峰（stutter peak），是引起额外峰第一个、也是最常见的原因，它主要出现在比其主峰（等位基因）短一个重复的位置，原因是在复制过程 TaqDNA 聚合酶的滑动。复制滑动在二核苷酸 STR 基因座较为多见，四、五核苷酸重复基因座则相对少见。在复合扩增中，由于扩增反应条件不是每个基因座的最佳条件，容易出现阴影带。

（2）3′端的额外 A 核苷酸添加

由于模板 DNA 量过多，或没有进行最后的 60℃保温处理，因 3′端的额外 A 核苷酸添加不完全会出现 N 和 N＋1 带，产生额外带。

（3）非特异人为带

在复合扩增体系，有时会出现非特异性片段，主要原因是多对引物发生的非特异结合。一般复合扩增的基因座越多，反应体系中引物对越多，非特异结合的倾向就越大。尤其是当样品 DNA 降解时更为严重，因为降解模板 DNA 互补序列变短。大多非特异带的峰面积比较低，从峰的形态看，大多表现为不正规的峰形，而且多不在基因座的等位基因范围内，很容易与等位基因峰区分开来。另外，模板 DNA 量过多，也会出现非特异带。

(4) 染色体的不正常

额外 DNA 带也可在一些由于像染色体移位、体细胞突变和三倍体等染色体不正常稀有的个体中遇到。杂合子个体在等位基因范围内有三条带，出现一个额外带。此各种情况下，物证斑痕样品与匹配的嫌疑人样品的 DNA 带型一定是相同的。因此当遇到这种少见的染色体不正常的个体识别案件时，所出现的额外 DNA 带可作为认定结论的一个依据。

(5) 分析软件的问题（不同荧光染料激发光的干扰峰）

标记 DNA 片段的染料在不同的波长发出荧光，仍存在一些染料吸收光谱的重叠，例如，蓝色标记 DNA 片段在绿色荧光吸收处也有荧光吸收。使用标准荧光染料预先制作的一个叫“matrix”的文件可以消除不同荧光之间的干扰。但分析过量扩增的样品（RUF 值超过 6000），饱和了‘matrix’，导致在等位基因峰下产生另一颜色的一个叫“pull - up”的干扰峰；如果选用不正确的‘Matrix’，即使荧光强度在正常范围内，也会产生一个这样的干扰峰（图 12 - 32）。因此在多色荧光标记分析中，通过峰的位置和峰形态比较，需要将这些“pull - up”峰同真正等位基因区分开来。另外也可以对这个基因座进行单个扩增分析来确定是否有“pull - up”峰。

当然，在银染系统或单色荧光标记分析中不存在这个问题。

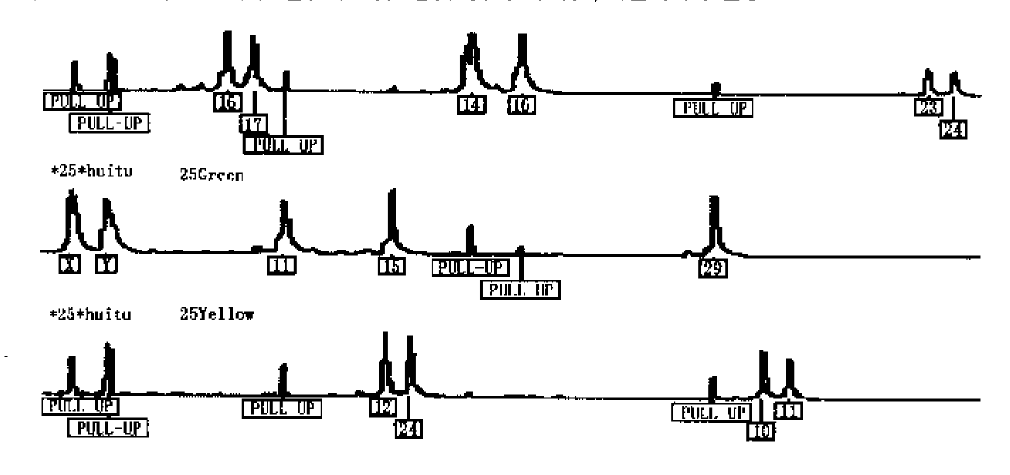


图 12 - 32 “pull up”峰

(6) 非产物荧光污染

荧光染料标记的 PCR 产物均采用光谱扫描系统如荧光扫描仪或 ABD PRISM 系统，利用的是可见光谱（500 ~ 600nn），因此一些具有荧光的物质如抗生素、维生素、多环芳香物、荧光素发生体和染料等在电泳图中出现与 DNA 荧光片段相同的峰，可能会干扰 DNA 分型，影响鉴定结果。Urquhart 及其同事（1993）发现 Chelex DNA 提取方法甚至有机提取法都不能去除抗生素、维生素、多环芳香物、荧光素发生体、各种染料等，以 Chelex 提取法提取的 DNA 为模板的 PCR 产物常伴有荧光污染峰。不过，干扰峰通常较宽，比较容易与荧光素染料标记 PCR 产物峰区分开来。

非产物荧光污染产生有以下四种可能情况：① 带色载体上的体液斑，在提取过程中染料被浸了下来；② 植物体上的体液斑，叶绿素可能与 DNA 同时提取出来；③ 病理状态下个体的血或组织样品，如铅中毒、某些形式的卟啉症，此时血中的卟啉浓度大幅度提

高；④ 年轻时曾接受过环组抗生素治疗的个体的牙或骨组织（成长的骨和牙齿吸收、蓄积了这些抗生素）。采用空白基质（载体）对照实验，或在没有合适空白基质做对照时，将未扩增的物质与 PCR 扩增产物一起电泳检测进行比较，确定是否有非产物污染峰。

（7）样品为混合样品

当分析样品为混合样品时，经常出现多带现象，这是造成额外带的最常见原因。混合样品的情况将在后面详细介绍。

（8）电泳条件不适当，FGA 基因座的带会移到 vWA 基因座区域，造成 vWA 基因座有三条带，而 FGA 基因座只有一条带，重新电泳检测可克服。

2. 等位基因峰面积的不对称

小卫星 VNTR 基因座各等位基因片段间差异较大，易产生等位基因的不平衡扩增，较小等位基因优先扩增，较大等位基因 PCR 产物量少或根本检测不到。在目前已知 STR 基因座中，等位基因片段均在 400bp 以下，二个等位基因之间差异小，也存在较低分子量的等位基因比较大分子量的等位基因具有高的扩增效率。但是，等位基因差异扩增不如小卫星 VNTR 那样显著，没有明显的优先扩增现象存在，因此峰面积基本相同。正常情况下，小峰面积为大峰面积的 70% 以上，因此 PE 公司将 70% 作为判定 2 个峰是否属于同一个体同一基因座的二个等位基因的标准。然而，下列情况可引起优先扩增，在应用中尤其是亲子鉴定中应引起注意。

（1）模板 DNA 极微量、DNA 降解、有抑制剂存在

模板 DNA 极微量或 DNA 降解或有抑制剂存在，在 PCR 开始几个循环，引物退火的随机差异，导致某一个等位基因优先扩增。另外，当一对等位基因的大小差异很显著时，也可引起一个等位基因的优先扩增。

如果一基因座的二个峰高比值小于 70%，而没有其他指标表明是混合样品，那么样品必须重新扩增以确定不平衡情况是否有重复性。不平衡扩增提示标本可能为混合样品。

（2）引物结合位置突变影响

模板 DNA 的引物结合位置出现突变，可能直接影响退火效率，尤其是突变发生在引物的 3′端，延伸可完全被阻断，使某个等位基因无扩增产物，成为无效等位基因（null allele），杂合子个体只有一个基因产物，成为所谓假纯合子的现象，如 MBP－B、D19S253 及 Profile plus 中 D8S1179 STR 系统均可出现上述无效等位基因，不过这种情况造成的假纯合子的等位基因峰高比真正纯合子的峰高低许多，可由此来确定是否有无效等位基因。如果突变发生在引物结合的其他位置，引物退火温度受到影响，导致信号降低。引物结合位置突变发生的机率极小（大约在 0.01～0.0001 基因座），单独一个个体不可能在二个基因座受到影响。重新设计引物序列，避开前一引物退火区域，可以克服无效等位基因的问题。

（3）样品为混合样品

在实际检验中，明显干扰峰面积平衡的常见的原因是混合样品，而不是等位基因的优先扩增。

当分析样品为混合样品时，混合样品组分的混合比例不同或虽混合的比例相同但拥有一个共同的等位基因，发生这种基因座内的优先扩增现象，等位基因峰面积明显不平衡。扩增分析多个基因座，综合分析整个图谱，确定是否为混合样品，有助于结果的解释。

3. 差异扩增

在二个或二个以上基因座复合扩增中，尤其是多个如荧光检测系统八个或九个基因座复合，有时并不是所有基因座都能得到扩增，也就是说部分基因座没有得到扩增产物。这种情况叫做差异扩增，多见于模板 DNA 量极微、模板 DNA 降解，或模板 DNA 中含有 PCR 反应的抑制剂如血液中的血红素、或其他色素。对于来自模板 DNA 的干扰因素，有些基因座不敏感，而有些则比较敏感，PCR 扩增则受到限制，产物的信号较弱或没有。一般的规律是等位基因片段相对大的基因座易受影响，例如 PE AmpFLSTR Profiler plus 系统中 D18S51、D7S820 基因座，其次是 FGA 基因座。在分析上述异常情况时要充分考虑上述各种影响因素，得出正确的结论。

12.5 STR 检测新技术

12.5.1 微芯片

1. 设计原理与过程

微芯片是通过光刻平版印刷和化学刻蚀方法在熔凝的石英晶片上形成带通道的结构，制成的小型电泳设备。40nm 厚的铬片喷溅在直径为 150mm 、厚 0.4mm 的熔凝石英晶片上。然后 1811 感光耐蚀膜被旋转包埋到厚度为 1.1μm 的晶片上，90℃烤 25min。通过选择性地暴露于 365nm 的紫外线下将光敏层刻成相应的图案。用 K_3 $(CN)_6$/NaOH 铬蚀除去暴露的铬，将晶片浸在 NH_4F/HF (1:1) 50℃刻蚀形成熔凝的石英。刻蚀的深度通过检测刻蚀时间和光度计测量控制。用丙酮处理除去感光耐蚀膜，剩下的铬溶解在 K_3 $(CN)_6$/NaOH 溶液中。在通道的末端用 CO_2 激光系统形成一直径为 75μm 的孔，它通过刻蚀晶片。第二个直径为 150mm 的熔凝石英晶片结合到刻蚀晶片上，封闭通道。将两个晶片浸在 50℃的 30% NH_4OH (wt/vol) /H_2O/30% H_2O_2 (wt/vol) 按 4:5:1 混合液中，然后，用过滤水彻底冲洗。这两个晶片连在一起，进行热结合。开始在 200℃下 2h 进行可逆性结合，然后在 1 000℃过夜进行永久性结合。这远在所用的熔凝石英底物软化温度下面，因此结合是在两个表面形成共价键，不会扩散到边缘分子。用晶片锯将单个微芯片从连接的晶片上切下来，形成高 3mm，内径 3mm 的 50μl 的槽。

如图 12-33 所示的带有直接分离通道的交叉结构仪器。通道 A、B，C 为 5mm 长，分离通道 D 长度为 30mm。芯片按反回归线被蚀刻成 45μm 深，顶端宽 100μm 的带半环交叉部分的通道。通道的交叉部相当于内径 70μm 的圆柱形毛细管。图 12-33b 描述的结构是用来研究通道折叠效果的一种常用技术，用来在微型仪器中制造压缩结构。通道 A，B，C 为 5mm 长，折叠的分离通道为 100 和 300mm 长。每个折叠的仪器的深度是 60μm，130μm 宽。折叠通道芯片内制成二种不同交叉几何图形 。

2. 包埋

微型仪器的内部通道表面用改进的 Hjerten 程序进行包埋。1.0mol/L NaOH 通过通道 10min，蚀刻 12h。然后，通道用过滤去离子水，0.1mol/L HCl，去离子水，MeOH 冲洗，每次 10min，在 He 6.0 条件下干燥。通道用 5.0ml 95% MeOH，0.5ml 10% 乙酸，1.0ml 3-三甲基 组成的溶液冲洗，12h 后，首先用 MeOH 冲洗，然后用去离子水冲洗，每次 10min。通道再在 He6.0 干燥。在最后步骤，70mg 丙烯酰胺溶解于 1.0ml 的分离缓冲液中，放置在

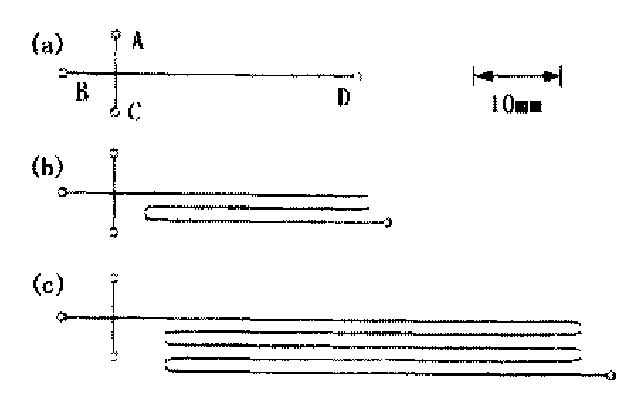


图 12-33 熔凝石英微芯片

He 6.0 中驱除氧。2h 后，加入 8μl 100%TEMED 和 20%过硫酸铵，简单混匀后，混合液通过注射器倒在蚀刻的通道中，在室温聚合 12h，用注射器除去过量聚合包埋溶液。通道已准备好，可以填充分离介质。

3. 分离介质

工作缓冲液为 1×TBE（90mmol/L Tris/64mmol/L 硼酸/2.5mmol/L EDTA，pH8.0），3.5mol/L 尿素，4%丙烯酰胺，30%甲酰胺。抽气 2h 后，加 2μl 10%TEMED，2μl 10%过硫酸铵，混合，将混合液灌到包埋的通道内。

4. 仪器

图 12-34 为微芯片基因分型仪器示意图。简单地说熔凝的石英微芯片安置在温度控制平台上，配有高压连接、激光诱导荧光检测仪。微芯片放在铝加热板上，加热板由温度控制器和一系列热偶联控制。

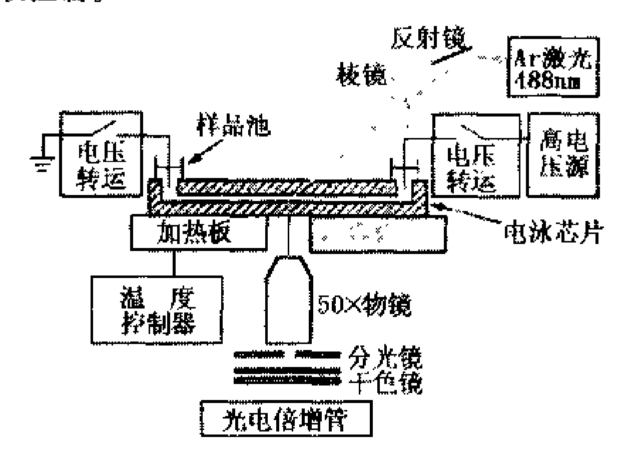


图 12-34 微芯片基因分型仪

5. 微芯片分离 STR 产物

为了快速分型，4 个通道都填充了聚丙烯酰胺介质。检测器放在离进样口 26mm 处，新灌制的芯片在 200V/cm 电压下 50℃预电泳 3min，，为了分离 CTTv 内部标准 Ladder，取 2μl Ladder 与 8μl 工作缓冲液混合。对于等位基因图谱，4μl PCR 扩增样品加到 2μl CTTv 的 Ladder 中，用含有 3.5mol/L 尿素和 30%甲酰胺的 2×TBE 稀释到 10μl，简单混合样品，

95℃变性 2min，迅速放置于冰水浴冷却，加到在通道 A 底部的微芯片样品瓶中，微芯片以 pinched cross injection 模式操作，在进样点出现三个离子流，确保样品离子到通道内部规定的体积。为了加样，在通道 A 和 C 加压 400V/cm，40V/cm 场强加到通道 B 和 D，防止样品进入这些通道。结果形成一个 100μm 长的稳定的进样活塞。进样量约 0.36nl，为了将样品注射到分离通道，给分离通道加场强为 200V/cm 的电压，通道 A 和 C 约 20V/cm，使样品很好地进入到分离通道 D 而没有从通道 A 和 C 的边上渗漏出来。用分离缓冲液冲洗样品池 A 三次后，更换样品，加样新的样品，进行几秒钟的预电泳。电泳分离样品，结果如图 12－35。

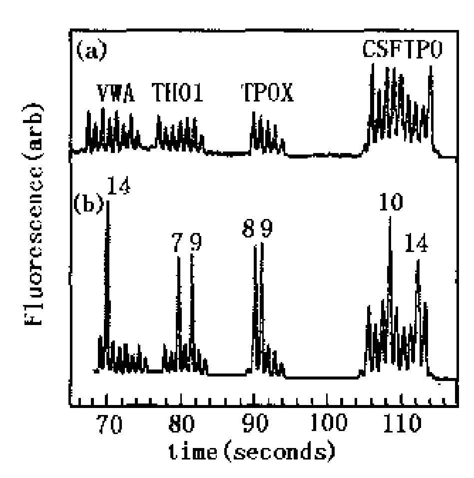


图 12－35　4 个基因座等位基因大小标准物的微芯片电泳轨迹图

(a) 芯片含有 45μm 深的通道，100μm 的样品进样器，26mm 长的分离通道。3.5mol/L 尿素和 30% 甲酰胺的 1×TBE 配制的 4% 线性聚丙烯酰胺在 50℃ 200 V/ cm 的场强下分离 (b) PCR 扩增样品与 CTTv 大小标准物混合的微芯片电泳图，图中所列的数字为样品的等位基因的分型，电泳条件同 (a)。

12.5.2 杂交排列分析 STR（STR 芯片）

除了凝胶电泳分析外，还可以用杂交模式分析 STR。美国 Nanogen 公司研发出另一种微芯片分析技术，可以进行 STR 分型。这种技术采用一种硅片，上面布有许多电极，每个电极为独立的实验点，可以根据需要连接上特定的核苷酸序列。电极上的核苷酸序列作为捕捉（capture）的探针，通过电杂交与 PCR 产物迅速杂交，这种杂交体系与另一叫报告（reporter）的探针杂交，形成稳定的杂交体。报告探针上带有荧光物，根据荧光点位置确定待测样品的基因分型。

图 12－36 是用杂交方式识别不同等位基因的原理示意图。对于 STR 基因座，芯片上包含该基因座所有已知的等位基因探针。STR 杂交分析涉及 2 个探针结合到 STR 重复和旁侧区（图 12－36A）。捕捉探针（capture probe）结合到芯片的实验点（图 12－36B），捕捉 PCR 扩增的 STR 等位基因，报告探针带有荧光染料，因此能够检测结合到捕捉探针的 STR 等位基因，重复区域任一边的的单一序列能够区别开拥有相同重复序列的不同 STR 基因座的等位基因。根据荧光信号出现位置，即根据捕捉探针的重复单位数目，确定样品基因型。

图 12－36A 是 THO1 基因座含有 6 个重复单位的靶序列，与含有 4，5 和 6 个重复单位的捕捉探针杂交，在最适当条件下，靶序列与每一位置的捕捉探针杂交几乎相当，但是，

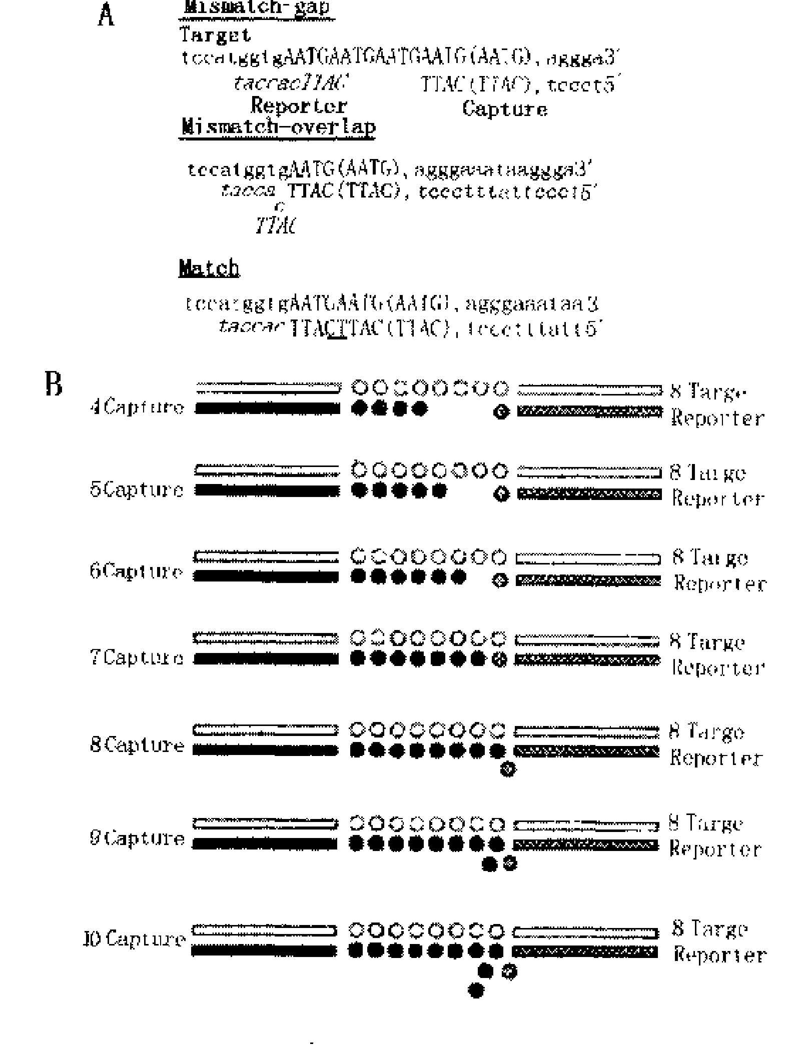


图 12-36 杂交进行 STR 基因分型的寡核苷酸系统

(A)：TH01 基因座的三种杂交复合物。不匹配-沟（mismatch-gap），在这种复合物中靶 DNA 中的重复数大于捕捉加报告探针寡核苷酸中的重复数；不匹配-重叠（mismatch-overlap），靶 DNA 中的重复数比报告和捕捉探针的重复数少；匹配（match），捕捉加报告探针中的重复数正好与靶 DNA 中的重复数相等。(B)：TH01 基因座等位基因 8 纯合子在芯片上可能的杂交模式。

每种情况下的捕捉探针与报告探针边界的部分关系是不同的，带 4 个重复的捕捉探针，杂交结果形成的捕捉/靶序列/报告探针杂交体在捕捉探针的 3′末端与报告探针 5′末端边界有一个孔隙，带 6 个重复的捕捉探针杂交，三者形成的杂交体在捕捉探针与探针之间有 4 个碱基的重叠，只有 5 个重复的捕捉探针形成的捕捉探针与报告探针的边界是毗连的，它们之间存在着碱基堆积力，在高强度下，只有这种杂交体最稳定，可以得到检测，其他几种杂交体由于没有碱基堆积力，在高强度下，报告探针被洗脱下来，没有检测信号。由此确定被检测样品的基因型。利用报告探针最大稳定性，结合报告探针和捕捉探针资料，可以鉴定 STR 等位基因的分型。

报告探针与靶序列互补碱基少于 12 个，杂交体可能非常不稳定。

稳定 DNA 杂交体的碱基堆积力受二个方面的因素影响。一是二个末端连接碱基的类型，二是碱基的烯丙基化修饰。例如 CSF1PO 基因座，选择捕捉探针和报告探针之间毗连碱基有 T:A 或 G:A，链内的 G:A 核苷酸的碱基堆积力比链间的 T:A 使螺旋稳定。捕捉探针与报告探针毗连末端以 G:A 方式存在的杂交体比与 T:A 系统的识别比率高（图 12－37A）。识别比率（discrimination ratio，DR）定义为从每个减去背景信号后的匹配/不匹配的比例。

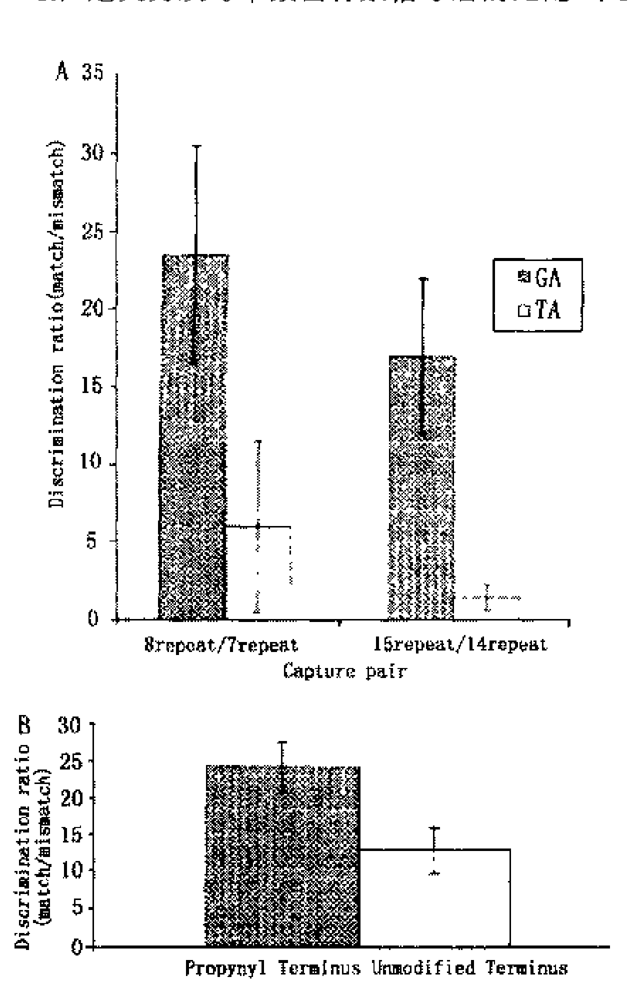


图 12－37 碱基堆积和杂交复合物的稳定性。

(A) CSF1PO 基因座的不同碱基堆积对的识别率的比较。芯片上点有重复数为 6、6.5、7、7.5、13、13.5、14 或 14.5 的捕捉探针寡核苷酸，与重复序列为 8 和 15 的靶 DNA 杂交。然后，二个 10nt 报告探针被动杂交到芯片上。一个探针的 5′端含有一个完整重复，另一个探针则含有半个重复单位。它们各自产生 G:A 或 T:A 的堆积作用。实验结果表明 G:A 报告:捕捉复合的识别率最高。(B) 带自然末端与带烯丙基化末端报告探针在 TH01 基因座的识别率比较。

碱基的烯丙基化可以增加碱基堆积，可能是增加碱基环结构的电子。偶极，π 键，偶极诱导的偶极机制形成环的电子增加，对核苷酸中垂直碱基堆积有重要作用。

利用杂交方式进行分型，要求杂交的高度特异性，在杂交过程中应用高强度条件。但是这需要非常长的杂交时间（3～16h）才能得到足够的信号。如果利用电子系统浓缩 DNA 样品，在高强度下进行电子主动杂交，可以在 2min 内完成。

许多 STR 基因座，除了重复单位数目变化外，还有由于插入、缺失、转换或置换形成的微变异体。这些微变异体，尤其是转换或置换造成的，很难用凝胶电泳方式检测出。但是，用杂交方法可以分析这些差异。以 TH01 为例来说明，TH01 基因座重复单位有一个缺失，即等位基因 9.3。一个生物素化的捕捉探针设置在电极上，其中 3 个为图 12－36 介绍的含 7，8 和 9 重复单位的，与 TH01 9.3 同源的靶 DNA（K562）在组氨酸缓冲液中加热变性，进行电杂交。带 BTR 标记的报告探针（专门为 TH019.3 变异体设计的）进行被动杂交，洗涤并检测，当与 9.3 等位基因杂交时，这个报告探针完全互补，垂直碱基堆积，但其他的等位基因的互补区从 10 个减少到 7 个，破坏了碱基堆积，不能得到检测信号。

12.5.3 MALDI－TOF 质谱

MALDI－TOF 利用离子进行实际时间的质量测定。图 12－38 是一分析原理示意图，液体 DNA 样品与过量的基质复合物结合，如 3－羟基甲基吡定酸（3－hydroxy－picolinic acid），然后这些样品滴到金属或石英板上、空气中晾干，DNA 和基质形成共结晶，然后把样品板放到质谱的真空环境中进行分析。快速激光脉冲激发离子化过程。DNA 分子周围的基质分子保护 DNA 在离子化过程中片段化。每个激光脉冲激发样品的离子化，然后在飞行管内分离离子。DNA 离子按其质量大小先后通过检测器，DNA 片段越短，越早到达检测器。移动固定激光束下的样品板顺序分析样品。市场上有样品板销售，一次可以点 384 个样品，在 1 小时内分析完每个样品板，分析时间取决于每个样品收集的激光点数和激光脉冲比。飞行时间质谱具有极高的通量，MALDI－TOF－MS 分析可以不使用等位基因分型标准物进行分析，因为此方法分析的是 DNA 分子的真实质量，使之比电泳中的相对大小测量更准确。但分析 STR 的 PCR 产物还存在一些问题，最大的问题是，只能分析片段较短的 DNA 分子，当 DNA 分子或样品盐浓度太大时，MALDI－TOF－MS 的分辨率和灵敏度下降。

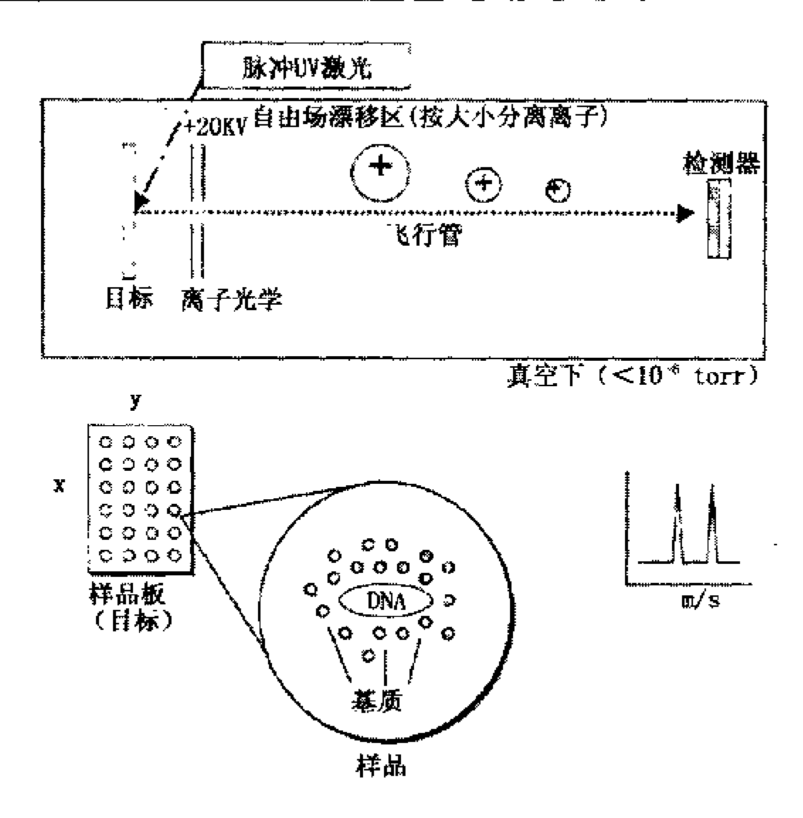


图 12-38 MALDI 时间飞行质谱分析原理示意图

DNA 样品与基质复合物混合，点在样品板上干燥，形成一个样品点。样品板进入质谱的真空环境，小部分样品在脉冲 UV 激光作用下离子化，形成的离子通过离子光束被加速。离子根据大小在通过飞行管时被分离开。当离子撞击检测器时每个离子的质/荷比被检测器记录下来。

第十三章　常染色体 STR 基因座

STR 基因座特性和分析方法等已在前一章介绍，本章着重介绍各个常染色体基因座特点、群体数据、引物序列与 PCR 反应条件。所有基因座的扩增产物检测，目前大都采用 4%、5%的变性聚丙烯酰胺凝胶电泳或毛细管电泳，各个基因座间没有大的差异，因此本章不再一一介绍。

13.1 D1S549 基因座

D1S549 基因座定位于人类 1 号染色体，核心序列为（GATA）n。在 100 名甘肃回族群体无关个体中发现 8 个等位基因，杂合度为 0.789，个人识别率为 0.910；在 109 名新疆维吾尔族群体无关个体中发现 7 个等位基因，杂合度为 0.85，个人识别率为 0.934。

1. 引物序列　5′ - CAA AGA GGA CAT GTG TTT GTG - 3′

　　5′ - TAC CAG CAA TCC CTA CTA TGG - 3′

2. PCR 反应　反应体积 20μl，DNA 2 ~ 40ng，1 × Taq 缓冲液，$MgCl_2$ 1.5mmol/L，Taq DNA 聚合酶 1u，dNTPs 150μmol/L，引物浓度为 0.25μmol/L。

3. 热循环参数　94℃变性 30s，58℃退火 30s，72℃延伸 40s，循环 30 个周期。

13.2 D3S1358 基因座

D3S1358 基因座定位于第 3 号染色体，结构上属于复合的 STR 系统，含有二个或多个不同的重复序列：[AGAN]，其中 N 为 T 或 C。等位基因重复区域的序列见表 13 - 1。已检出 12 个等位基因，等位基因片段长度范围为 103 ~ 147bp，杂合度为 0.79，Dp 值为 0.92。群体等位基因频率见表 13 - 2。

表 13 - 1　D3S1358 基因座序列

等位基因	重复区域序列	片段大小（bp）
13	[AGAT] 9 [AGAC] 2 [AGAT] 2	119
14	[AGAT] 10 [AGAC] 2 [AGAT] 2	123
15	[AGAT] 11 [AGAC] 2 [AGAT] 2	127
16	[AGAT] 12 [AGAC] 2 [AGAT] 2	131
16′	[AGAT] 11 [AGAC] 3 [AGAT] 2	131
17	[AGAT] 13 [AGAC] 2 [AGAT] 2	135
17′	[AGAT] 12 [AGAC] 3 [AGAT] 2	135
18	[AGAT] 13 [AGAC] 2 [AGAT] 2	139
19	[AGAT] 14 [AGAC] 3 [AGAT] 2	143
20	[AGAT] 15 [AGAC] 3 [AGAT] 2	147

表 13-2 D3S1358 基因座等位基因频率

等位基因	频				率	
	中国汉族	中国汉族	中国汉族(北京)	中国汉族(武汉)	美国白人	美国黑人
9						0.0026
11					0.0025	0.0026
12	0.0022	0.0017			0.0025	0.0051
13		0.0033	0.0017	0.0026	0.0050	
14	0.0459	0.0417	0.0433	0.0464	0.1125	0.1180
15	0.4039	0.3433	0.3700	0.3608	0.2825	0.2795
15.2						0.0026
16	0.2751	0.3333	0.3150	0.3325	0.2225	0.3231
17	0.1989	0.2017	0.1917	0.2216	0.2225	0.2180
18	0.0655	0.0717	0.0667	0.0309	0.1450	0.0462
19	0.0087	0.0033	0.0117	0.0052	0.0050	0.0026
杂合度			0.7219	0.7423		
DP			0.9343	0.8518		

1. 引物序列 5' - ACT GCA GTC CAA TCT GGG T - 3'

5' - ATG AAA TCA ACA GAG GCT TG - 3'(M.Gene, 1998)

2. PCR 反应 在 50μl 反应体系中含有模板 DNA 10ng, 1× PCR 缓冲液(50mmol/L KCl, 10mmol/L Tris - HCl, 1.0mmol/L $MgCl_2$), dNTPs 200μmol/L 和 Taq DNA 聚合酶 1.25u, 每种引物 0.25μmol/L。

3. 热循环参数 94℃ 5min 预变性, 94℃ 1min, 57℃ 45s, 72℃ 45s, 30 个循环, 接着 72℃ 10min。

13.3 D3S1359 基因座

D3S1359 基因座(GenBank No.AA306290)定位于第 3 号染色体, 结构上属于复合的 STR 系统, 含有二个或多个不同的重复序列: TCTA 和 TGTA(表 13-3)。已检出 17 个等位基因, 等位基因片段大小为 196~265bp。

表 13-3 D3S1359 基因座序列

等位基因	片段大小	序列结构
11	196	$[TCTA]_{11}$
12	200	$[TCTA]_{12}$
13	204	$[TCTA]_{13}$
14	208	$[TCTA]_{14}$
15	212	$[TCTA]_{15}$
16	216	$[TCTA]_{16}$
17.3	223	$[TCTA]_6$ $[TGTA]_1$ $[TCTA]_1$ TCA $[TCTA]_9$

续表

18	224	[TCTA]$_{18}$
18.3	227	[TCTA]$_6$ [TGTA]$_1$ [TCTA]$_1$ TCA [TCTA]$_{10}$
19.3	231	[TCTA]$_7$ [TGTA]$_1$ [TCTA]$_1$ TCA [TCTA]$_9$
20.3	235	TC [TCTA]$_6$ [TGTA]$_1$ [TCTA]$_1$ TCA [TCTA]$_{12}$
21.3	239	[TCTA]$_7$ [TGTA]$_1$ [TCTA]$_1$ TCA [TCTA]$_{12}$
21.3b	239	[TCTA]$_9$ [TGTA]$_1$ [TCTA]$_1$ TCA [TCTA]$_{10}$
22.3	243	[TCTA]$_7$ [TGTA]$_1$ [TCTA]$_1$ TCA [TCTA]$_{13}$
22.3b	243	[TCTA]$_9$ [TGTA]$_1$ [TCTA]$_1$ TCA [TCTA]$_{11}$
23.3	247	[TCTA]$_9$ [TGTA]$_1$ [TCTA]$_1$ TCA [TCTA]$_{12}$
24.3	251	[TCTA]$_9$ [TGTA]$_1$ [TCTA]$_1$ TCA [TCTA]$_{13}$
25.3	255	[TCTA]$_9$ [TGTA]$_1$ [TCTA]$_1$ TCA [TCTA]$_{14}$

1. 引物序列　5′ - GTT GCC TCT GAC ATG GCT TT - 3′

5′ - ATG CTA AGT GCT AAG TCA ACT - 3′

2. PCR 反应　50μl 反应体系，含有 10× PCR 缓冲液（100mmol/L Tris - HCl，pH 9.0，500mmol/L KCl，15mmol/L $MgCl_2$，1mg/ml 白明胶）2μl，0.5μmol/L dNTPs 1.6μl，每种引物 25pmol，Taq DNA 聚合酶 2.5u，模板 DNA 15ng。

3. 热循环参数　首先 94℃预变性 5min，然后经 94℃ 30s，60 ℃45s，72℃30s，30 次循环。

13.4 D3S1744 基因座

D3S1744 基因座定位于人类染色体 3q24。在德国人群已检出 9 个等位基因，等位基因片段长度大小范围为 150 ~ 182bp，杂合度为 0.814，非父排除率为 0.632，多态信息量为 0.787。

1. 引物序列　5′ - TTT AAG CGG AAC GAA GTG TG - 3′

5′ - CTG GCC CCA TCT CTC TCT AT - 3′

2. PCR 反应　在 10μl 反应体系中，含有 10mmol/L Tris - HCl，pH 8.3，50mmol/L KCl，1.5mmol/L $MgCl_2$，200μmmol/L dNTPs，每种引物 1pmol，Taq DNA 聚合酶 0.3u，模板 DNA 30ng。

3. 热循环参数　94℃ 30s，55℃ 75s，72℃延伸 15s，27 个循环，最后 72℃保温 6min。

13.5 D5S818 基因座

D5S818 基因座定位于人染色体 5q21 - 31，重复序列为 ACAT，等位基因片段大小为 135 ~ 171bp，等位基因频率见表 13 - 4。

表 13-4　D5S818 基因座等位基因频率

等位基因	频率				
	中国汉族	中国汉族（北京）	中国汉族	美国白人	美国黑人
6	0.0044				
7	0.0133	0.0200	0.0183	0.0025	0.0026
8	0.0067		0.0017	0.0050	0.0513
9	0.0844	0.0800	0.080	0.0225	0.0205
10	0.1756	0.1467	0.2000	0.0675	0.0744
11	0.3333	0.4067	0.3300	0.3925	0.2539
12	0.2467	0.2050	0.2100	0.3325	0.3256
13	0.1244	0.1267	0.1450	0.1650	0.2487
14	0.0089	0.0150	0.0067	0.010	0.0205
15	0.0022		0.0083		0.0026
16				0.0025	0
杂合度		0.7493	0.7667		
DP		0.9573	0.9284		

1. 引物序列　5′-GGG TGA TTT TCC TCT TTG GT-3′

　　　　　　5′-TGA TTC CAA TCA TAG CCA CA-3′

2. PCR 反应　在 10μl 反应体系中，含有 10mmol/L Tris-HCl，pH 8.3，50mmol/L KCl，1.5mmol/L $MgCl_2$，200μmmol/L dNTPs，1pmol 每种引物，Taq DNA 聚合酶 0.3u，模板 DNA 30ng。

3. PCR 热循环参数　94℃ 30s，55℃ 75s，72℃延伸 15s，27 个循环，最后 72℃保温 6min。

13.6 D6S366 基因座

D6S366 基因座定位于人类染色体 6q21-qter，为三核苷酸重复单位（ATT），已检出 8 个等位基因，等位基因片段长度范围 138～162bp 频率分布见表 13-5。杂合度为 0.72，个人识别率为 0.877。

表 13-5　D6S366 基因座等位基因分布频率

等位基因	频率	
	中国汉族（武汉）*	中国汉族（成都）
13	0.03	0.0202
14		
15	0.015	0.0606
16	0.075	0.3737
17	0.2975	0.4747
18	0.415	0.0556
19	0.1275	0.0152
	0.04	
杂合度		0.68
DP		0.78

* 按等位基因片段大小命名，最小的为等位基因 1，与左边等位基因命名可能不同

1. 引物序列 5′ - AGA GGT TAC AGT GAG CCG AGA TTG - 3′

5′ - GAA GTC CTA ACA GAA TGG AAG GTC C - 3′

2. PCR 反应 20μl 反应体系中，含有 10 × PCR 缓冲液（100mmol/L Tris - HCl，pH 9.0，500mmol/L KCl，15mmol/L$MgCl_2$，1mg/ml 白明胶）2μl，2.5mmol/L dNTPs 1.6μl，每种引物（4μmol/L）2μl，Taq DNA 聚合酶 1μl（0.5u），蒸馏水 9.4μl，模板 DNA 2μl。

3. 热循环参数 96℃2min，然后经 94℃ 1min，60 ℃1min，70℃1.5min，10 次循环，90℃ 1min，60℃ 1min，70℃ 1.5min，20 个循环，最后 60℃保温 30min。

13.7 D7S809 基因座

D7S809 基因座又叫 wg1g9 基因座，核心序列为 AGGM（M = A 或 C），在日本人中共发现 13 个等位基因，片段长度范围为 241 ~ 28bp，等位基因 9 频率最高，杂合度为 0.86。

1. 引物序列 5′ - GAT CGT GCC ACT GTA CC - 3′

5′ - CGC GTG TCT AAC CTC ATG GC - 3′

2. PCR 反应 在 10μl 反应体系中，含有 1.5pmol 引物 wg1g9B，10pmol 引物 wg1g9A，1 × PCR 缓冲液（标准缓冲液），0.5u Taq DNA 聚合酶，模板 DNA 10ng。

3. 热循环参数 96℃ 1min，62℃ 1min，70℃ 2min，循环 18 次。

13.8 D7S817 基因座

D7S817 基因座定位于人类第 7 号染色体，其核心序列分别为（ATCT）n。在 107 名成都汉族群体中发现 9 个等位基因，杂合度为 0.738，个人识别率为 0.931；在 101 名东乡族群体中发现 8 个等位基因，杂合度为 0.752，个人识别率为 0.917；在 106 名回族群体中发现 6 个等位基因，杂合度为 0.755，个人识别率为 0.895；在 104 名维吾尔族群体中发现 7 个等位基因，杂合度为 0.779，个人识别率为 0.930。

1. 引物序列 5′ - TTG GGA CCT CTT ATT TTC CA - 3′

5′ - GGC TTC TGC AGA GAA ACA GA - 3′

2. PCR 反应 2 ~ 40ng 基因组 DNA，1 × Taq 缓冲液，1.5mmol/L $MgCl_2$，1uTaq DNA 聚合酶，150μmol/L dNTPs，引物各 0.25μmol/L，反应体积 20μl。

3. 热循环参数 94℃35s，56℃45s，72℃55s，32 个循环。

13.9 D7S820 基因座

D7S820 基因座定位于人类染色体 7q11.21 - 22，核心序列为 GATA，已检出 9 个等位基因（表 13 - 6），核心序列重复次数为 6 ~ 14，等位基因片段大小范围为 215 ~ 247bp。

表 13-6 D7S820 基因座等位基因频率

等位基因	频率									
	中国汉族	中国汉族	中国汉族（北京）	中国汉族（北京）	中国汉族（河北）	中国汉族（太原）	中国回族（宁夏）	中国壮族（广西）	美国白人	美国黑人
5	0.0017	0	0				0		0.002	0.002
6										
7	0.0017	0	0		0.0028		0.0049	0.0074	0.010	0.012
8	0.1233	0.1287	0.2026	0.2030	0.1397	0.1238	0.1165	0.1324	0.155	0.179
9	0.0567	0.8042	0.1718	0.1716	0.0950	0.0842	0.0825	0.0662	0.152	0.084
10	0.1800	0.1535	0.1718	0.1716	0.1508	0.1634	0.1942	0.1765	0.295	0.351
11	0.3738	0.2970	0.2410	0.2410	0.3520	0.3069	0.3204	0.3971	0.195	0.235
12	0.2200	0.2624	0.1846	0.1846	0.2151	0.2574	0.2524	0.1618	0.121	0.112
13	0.0367	0.0693	0.0256	0.0256	0.0307	0.0594	0.0194	0.0515	0.057	0.019
14	0.0017	0.005	0.0026	0.0033	0.0028	0.0050	0.0097	0.0074	0.012	0.007
H	0.7641	0.791		0.8092		0.79	0.8350		0.80	0.76
DP	0.9026	0.9260		0.9336		0.9238	0.9150	0.9032		

1. 引物序列 5′-TGT CAT AGT TTA GAA CGA ACT AAC G-3′

5′-CTG AGG TAT CAA AAA CTC AGA GG-3′

2. PCR 反应 在 10μl 反应体系中，含有 10mmol/L Tris-HCl，pH 8.3，50mmol/L KCl，1.5mmol/L $MgCl_2$，200μmmol/L dNTPs，每种引物 1pmol，Taq DNA 聚合酶 0.3u，模板 DNA 30ng。

3. 热循环参数 94℃ 30s，55℃ 75s，72℃15s，27 个循环，最后 72℃保温 6min。

13.10 D8S384 基因座

D8S384 基因座定位于人类第 8 号染色体上，是一个由（TAGA）n 构成的四核苷酸重复序列。它最初作为人类基因组序列标签（sequence tagged site，STS），由 Utah 大学人类遗传标记工作组分离获得，命名为人类染色体 8STS UT1205（GenBank，L18535）。已检出 8 个等位基因（表 13-7），等位基因片段大小范围为 204～232bp。

表 13-7 D8S384 基因座等位基因频率

等位基因	频率							
	中国汉族（广东）	中国汉族（吉林）	中国汉族（成都）	中国壮族（广西）	中国白族（云南）	中国藏族（西藏）	中国蒙古族（内蒙古）	日本人
5	0.0050	0.0098	0.0158	0	0.0052	0	0	0
6	0.0202	0.0294	0.0053	0.0048	0	0.0109	0.0474	0.0084
7	0.1616	0.1128	0.1526	0.2143	0.1510	0.2119	0.1579	0.1639
8	0.4192	0.5098	0.4368	0.4095	0.4740	0.4620	0.2842	0.3740
9	0.2920	0.2353	0.3105	0.2810	0.2917	0.2337	0.2684	0.3277
10	0.0859	0.0882	0.0737	0.0714	0.0677	0.0706	0.1737	0.0924
11	0.0152	0.0147	0.0053	0.0190	0.0104	0.0109	0.0474	0.0336
12	0	0.	0	0	0	0	0.0210	0
杂合度	0.727	0.676	0.726	0.733	0.563	0.620	0.726	0.756
DP	0.864	0.844	0.846	0.859	0.831	0.848	0.923	0.871

1. 引物序列　5′ – TTT CTC AGT ATT CTA CAC AGG – 3′

5′ – GTT CCT GTC TTC TTC TAC A G – 3′

2. PCR 反应　反应体积为 37.5μl，含有 2 ~ 4ng 基因组 DNA，1 × Taq 缓冲液，1.5mmol/L $MgCl_2$，160μmol/L dNTPs，1.5u Taq DNA 聚合酶，每种引物各 0.25μmol/L。

3. 热循环参数　94℃1min，60℃ 1min，72℃1min，循环 30 次。

13.11 D8S1132 基因座

D8S1132 基因座定位于人类第 8 号染色体上，由数种重复序列组成，是较为复杂的 STR 基因座。重复序列结构为 $(TCTA)_{5-11}$ CA $(TCTA)_{9-13}$ $(TCTG)_{0-1}$ (TCTA)。片段长度规律地按 4bp 递增，等位基因片段大小为 134 ~ 170bp。等位基因的序列组成见表 13 – 8。

表 13 – 8　D8S1132 等位基因命名与序列组成

等位基因命名	片段长度	序列组成
16	134bp	$(TCTA)_5$TCA $(TCTA)_9$TCTG TCTA
17	138 bp	$(TCTA)_5$TCA $(TCTA)_{10}$TCTG TCTA
		$(TCTA)_6$TCA $(TCTA)_{10}$ TCTA
18	142 bp	$(TCTA)_5$TCA $(TCTA)_{11}$TCTG TCTA
		(TCTA) 6TCA $(TCTA)_{11}$ TCTA
19	146 bp	$(TCTA)_8$TCA $(TCTA)_9$TCTG TCTA
		$(TCTA)_6$TCA $(TCTA)_{10}$ TCTA
20	150 bp	$(TCTA)_5$TCA $(TCTA)_{13}$TCTG TCTA
		$(TCTA)_6$TCA $(TCTA)_{13}$ TCTA
		$(TCTA)_9$TCA $(TCTA)_{10}$ TC
21	154 bp	$(TCTA)_{10}$TCA $(TCTA)_{10}$ TCTA
		$(TCTA)_9$TCA $(TCTA)_{10}$TCTG TCTA
		$(TCTA)_8$TCA $(TCTA)_{11}$TCTG TCTA
22	158 bp	$(TCTA)_{10}$TCA $(TCTA)_{11}$ TCTA
		$(TCTA)_9$TCA $(TCTA)_{11}$TCTG TCTA
		$(TCTA)_5$TCA $(TCTA)_{10}$TCTG TCTA
23	162 bp	$(TCTA)_{11}$TCA $(TCTA)_{11}$ TCTA
		$(TCTA)_{11}$TCA $(TCTA)_{10}$TCTG TCTA
24	166 bp	$(TCTA)_{11}$TCA $(TCTA)_{12}$ TCTA
25	170 bp	$(TCTA)_{11}$TCA $(TCTA)_{13}$ TCTA

1. 引物序列　5′ – GGC TAG GAA AGG TTA GTG GC – 3′

5′ – CCC TCT CTC TTT CGA GCA AT – 3′

2. PCR 反应　25μl 反应体系，含有 10mmol/L Tris – HCl (pH 9.0)，50mmol/L KCl，1.5mmol/L $MgCl_2$，0.001% 白明胶，0.1% Triton – X 100，10μg BSA，100 μmol/L dNTPs，每种引物 0.2μmol/L，Taq DNA 聚合酶 1u，Chelex – 100 提取的模板 DNA 1 ~ 2μl。

3. 热循环参数　96℃ 2min 后，94℃ 1min，61℃ 1min，72℃ 1min，30 个循环。

13.12 D8S1179 基因座

D8S1179 基因座定位于人类染色体 8，核心序列为 TCTR，其中 R 为 G 或 A（早期被误认为 D6S502 基因座）。已查出 12 个等位基因，一些群体等位基因频率见表 13-9。等位基因片段长度范围为 162~206bp。小片段等位基因（162~178bp）只拥有简单四核苷酸 TCTA 重复序列，而较大等位基因（182~202bp）含有 TCTA 和 TCTG 二个四核苷酸组成的复合重复单位，测序证明等位基因 5′或 3′端没有序列变异，各等位基因重复单位序列如下表 13-10。

表 13-9 D8S1179 基因座等位基因频率

等位基因	频率						
	中国汉族	中国汉族（北京）	中国汉族（成都）	中国汉族（重庆）	中国汉族	美国白人	美国黑人
8		0.0017	0.000		0.0066	0.0175	0.000
9	0.0017	0.0017			0.0000	0.010	0.0051
10	0.1200	0.0883		0.1130	0.0881	0.0800	0.0308
11	0.0900	0.0567	0.0644	0.0830	0.0639	0.0625	0.0436
12	0.1050	0.1033	0.1237	0.1330	0.1278	0.1425	0.1051
13	0.2100	0.2683	0.1089	0.1840	0.2313	0.3475	0.1974
14	0.1967	0.2500	0.1386	0.1930	0.2048	0.1875	0.3333
15	0.1833	0.1733	0.2079	0.2060	0.1674	0.1300	0.2103
16	0.0683	0.0550	0.1931	0.0780	0.0969	0.0200	0.0641
17	0.0217	0.0017	0.1040	0.0080	0.0132	0.0025	0.0103
18	0.0033		0.0396	0.0020			
19			0.0198				
杂合度	0.8367	0.8121	0.91	0.8310	0.8106		
DP	0.9642	0.9666	0.94	0.9540	0.9562		

表 13-10 D8S1179 等位基因重复序列

等位基因	重复序列
8	$(TCTA)_8$
9	$(TCTA)_9$
10	$(TCTA)_{10}$
11	$(TCTA)_{11}$
12	$(TCTA)_{12}$
13	TCTATCTG $(TCTA)_{11}$
14	$(TCTA)_2$TCTC $(TCTA)_{11}$
15	$(TCTA)_2$TCTG $(TCTA)_{12}$
16	$(TCTA)_2$TCTG $(TCTA)_{13}$
17	$(TCTA)_2$ (TCTG) 2 $(TCTA)_{13}$
18	$(TCTA)_2$TCTG $(TCTA)_{15}$

1. 引物序列 5′-TTT TTC TAT TTC ATG TGT ACA TTC G-3′

5′-CGT AGC TAT AAT TAG TTC ATT TTC -3′

2. PCR 反应 20μl 反应体系，含有 10mmol/L Tris-HCl (pH 9.0)，50mmol/L KCl，1.5mmol/L $MgCl_2$，0.001% 白明胶，0.1% Triton-X 100，10μg BSA，200 μmol/L dNTPs，每种引物 12 pmol，Taq DNA 聚合酶 0.5u，模板 DNA 5ng。

3. 热循环参数 95℃ 1min，59℃ 1min，72℃ 1min 20s，32 个循环。

在实际案件中材料分型时，偶有发现 PE 公司 Profiler plus 试剂盒中 D8S1179 基因座出现等位基因丢失，形成一个无效基因的现象，推测原因可能是模板 DNA 的引物结合部位发生了点突变，影响了 PCR 中的退火过程。分析中根据峰面积判断或重新设计引物序列，可以消除此现象。

13.13 D10S1213 基因座

D10S1213 基因座位于人类第 10 号染色体上，核心序列为 CTTC，频率见表 13-11。

表 13-11 D10S1213 基因座等位基因频率

等位基因	频率	
	中国汉族（成都）	中国东乡族（甘肃）
6	0.090	0.010
8	0.045	0.040
9	0.527	0.364
10	0.195	0.308
11	0.136	0.172
12	0.036	0.010
13	0.027	0.071
14	0.018	0.020
15	0.005	0.005
杂合度	0.664	0.657
DP	0.836	0.882
CE	0.374	0.364

1. 引物序列 5′-CTA CAG ACA AAC CCA CTG GC-3′

5′-AGA GAG AAG GAA AGA AAC AAA GG-3′

2. PCR 反应 反应体系 20μl，模板 DNA2～4ng，引物 0.1μmol/L，dNTP 200μmol/L，Taq DNA 酶 1.25u 和 $MgCl_2$ 1.5mmol/L。

3. 热循环参数 94℃ 35s，58℃ 30s，72℃ 55s，循环 31 个周期，最后 72℃延伸 5min。

13.14 D10S1432 基因座

D10S1432 基因座位于人类第 10 号染色体上，核心序列为 ATTC 或 ATCT。基因频率分布见 13-12。

表 13-12 D10S1432 基因座等位基因频率

等位基因	频 率	
	中国汉族（成都）	中国东乡族（甘肃）
12	0.050	0.051
13	0.177	0.172
14	0.332	0.389
15	0.414	0.359
16	0.027	0.030
杂合度	0.664	0.737
DP	0.827	0.820
CE	0.374	0.488

1. 引物序列 5′-CAG TGG ACA CTA AAC ACA ATC C-3′

5′-TAG ATT ATC TAA ATG GTG GAT TTC C-3′

2. PCR 反应 反应体系 20μl，模板 DNA 2～4ng，引物 0.1μmol/L，dNTPs 200μmol/L，Taq DNA 酶 1.25u 和 $MgCl_2$ 1.5mmol/L。

3. 热循环参数 94℃ 35s，58℃ 30s，72℃ 55s，31 个循环，最后 72℃5min。

13.15 D10S2325 基因座

D10S2325 基因座定位于人类 10 号染色体，核心序列为 TCTTA，已检出 11 个等位基因，等位基因片段长度为 113～163bp。

1. 引物序列 5′-CTC ACG AAA GAA GCC TTC TG-3′

5′-GAG CTG AGA GAT CAC GCA CT-3′

2. PCR 反应 在 10μl 反应体系中，含有 10mmol/L Tris-HCl（pH 8.3），50mmol/L KCl，1.5mmol/L $MgCl_2$，200μmmol/L dNTPs，每种引物 1pmol，Taq DNA 聚合酶 0.3u，模板 DNA 30ng。

3. 热循环参数 94℃ 30s，55℃ 75s，72℃延伸 15s，27 个循环，最后 72℃保温 6min。

13.16 D12S375 基因座

D12S375 基因座定位于 12 号染色体，核心序列为 TATC，在中国汉族中共发现 5 个等位基因（频率分布见表 13-13），等位基因片段大小为 184～200bp。

表 13-13 D12S375 基因座等位基因频率

等位基因	频 率
	中国汉族
13	0.1905
14	0.0810
15	0.4714
16	0.2000
17	0.0571
杂合度	0.69
DP	0.82

1. 引物序列　5′ – TTG TTG AGG GTC TTT CTC CA – 3′

　　　　　　5′ – TCT TCT TAT TTG GAA TAA CCC – 3′

2. PCR 反应　20μl 反应体系，含有：模板 DNA 2 ~ 40ng，Taq DNA 聚合酶 0.25u，dNTPs 3.5μl（含 dNTPs 0.6μmol），1.25μmol/L 引物，10 × PCR 缓冲液 2μl，无菌水 11.75μl。

3. 热循环参数　94℃5min 后，94℃ 45s，58℃ 30s，72℃ 30s，30 个循环，72℃保持 5min。

13.17 D12S391 基因座

D12S391 基因座定位于人类 12 号染色体，核心序列为 (AGAT)$_N$ (AGAC)$_M$ (AGAT)$_L$，14 ~ 28 个重复，等位基因片段长度为 205 ~ 265bp。等位基因重复区域序列结构见表 13 – 14。等位基因频率见表 13 – 15。

表 13 – 14　D12S391 基因座序列结构

等位基因	重复结构
15	…… (AGAT)$_8$ (AGAC)$_6$ (AGAT)$_1$
16	…… (AGAT)$_8$ (AGAC)$_7$ (AGAT)$_1$
等位基因	重复结构
	…… (AGAT)$_9$ (AGAC)$_6$ (AGAT)$_1$
17	…… (AGAT)$_{10}$ (AGAC)$_6$ (AGAT)$_1$
	…… (AGAT)$_9$ (AGAC)$_7$ (AGAT)$_1$
	…… (AGAT)$_{11}$ (AGAC)$_5$ (AGAT)$_1$
18	…… (AGAT)$_{11}$ (AGAC)$_6$ (AGAT)$_1$
	…… (AGAT)$_{10}$ (AGAC)$_7$ (AGAT)$_1$
19	…… (AGAT)$_{12}$ (AGAC)$_6$ (AGAT)$_1$
	…… (AGAT)$_{11}$ (AGAC)$_7$ (AGAT)$_1$
	…… (AGAT)$_{10}$ (AGAC)$_8$ (AGAT)$_1$
20	…… (AGAT)$_{13}$ (AGAC)$_6$ (AGAT)$_1$
	…… (AGAT)$_{12}$ (AGAC)$_7$ (AGAT)$_1$
	…… (AGAT)$_{12}$ (AGAC)$_8$
	…… (AGAT)$_{13}$ (AGAC)$_7$
21	…… (AGAT)$_{12}$ (AGAC)$_8$ (AGAT)$_1$
	…… (AGAT)$_{12}$ (AGAC)$_9$
	…… (AGAT)$_{13}$ (AGAC)$_7$ (AGAT)$_1$
	…… (AGAT)$_{11}$ (AGAC)$_{10}$
	…… (AGAT)$_{14}$ (AGAC)$_7$
22	…… (AGAT)$_{13}$ (AGAC)$_8$ (AGAT)$_1$
	…… (AGAT)$_{13}$ (AGAC)$_9$
	…… (AGAT)$_{15}$ (AGAC)$_7$ (AGAT)$_1$
23	…… (AGAT)$_{14}$ (AGAC)$_8$ (AGAT)$_1$
	…… (AGAT)$_{14}$ (AGAC)$_9$
	…… (AGAT)$_{15}$ (AGAC) 7 (AGAT)$_1$
	…… (AGAT)$_1$ (AGAC)$_{13}$ (AGAT)$_1$
24	…… (AGAT)$_{15}$ (AGAC)$_8$ (AGAT)$_1$
	…… (AGAT)$_{14}$ (AGAC)$_9$ (AGAT)$_1$
	…… (AGAT)$_{13}$ (AGAC)$_{10}$ (AGAT)$_1$
	…… (AGAT)$_{15}$ (AGAC)$_9$
	…… (AGGT)$_1$(AGAT)$_{14}$(AGAC)$_8$(AGAT)$_1$
25	…… (AGAT)$_{16}$ (AGAC)$_8$ (AGAT)$_1$
	…… (AGAT)$_{15}$ (AGAC)$_9$ (AGAT)$_1$
	…… (AGAT)$_{16}$ (AGAC)$_9$
	…… (AGGT)$_1$(AGAT)$_{14}$(AGAC)$_9$(AGAT)$_1$
	…… (AGGT)$_1$(AGAT)$_{15}$(AGAC)8(AGAT)$_1$
26	…… (AGAT)$_{17}$ (AGAC)$_8$ (AGAT)$_1$
27	…… (AGGT)1 (AGAT)16(AGAC)9(AGAT)1
28	…… (AGAT)$_{18}$ (AGAC)$_9$ (AGAT)$_1$

表 13－15 D12S391 基因座等位基因频率

等位基因	频率				
	中国汉族	中国汉族（武汉）	日本人	白人	德国人
14					
15	0.004	0.0096		0.063	0.028
16	0.009	0.0024	0.01	0.025	0.02
17	0.052	0.0649	0.09	0.127	0.117
17.3				0.022	
18	0.191	0.2139	0.295	0.196	0.186
18.3				0.009	
19	0.255	0.2019	0.22	0.111	0.137
19.3				0.009	
20	0.189	0.1659	0.135	0.098	0.113
21	0.128	0.1394	0.075	0.123	0.137
22	0.086	0.1154	0.04	0.098	0.105
22.3				0.003	
23	0.05	0.0409	0.035	0.076	0.093
24	0.025	0.0361	0.01	0.025	0.057
25	0.009	0.0096	0.045	0.009	0.004
26	0.002		0.02	0.003	0.004
27			0.005		
28			0.005		
杂合度	0.847	0.9842	0.85	0.85	0.895
DP	0.95	0.9523	0.95	0.97	0.97
PIC		0.8273			

1. 引物序列 5′－ AAC AGG ATC AAT GGA TGC AT－3′

5′－TGG CTT TTA GAC CTG GAC TG－3′

2.PCR 反应 在 50μl 反应体系中，含有 50ng DNA，10mmol/L Tris－HCl，pH 9.0，500mmol/L KCl，15mmol/L $MgCl_2$，1mg/ml 白明胶 2μl，200μmol/L dNTPs，每种引物 0.5μmol/L，Taq DNA 聚合酶 1u。

3. 热循环参数 94℃10min 后，94℃45s，56℃1min，72℃7min，循环 30 次。

13.18 D13S317 基因座

D13S317 基因座定位于人类染色体 13q22－q31，核心序列为 AGAT，已检出 9 个等位基因，核心序列重复次数为 7～15，等位基因片段长度大小范围为 165～197bp，频率见表 13－16。

表 13-16　D7S820 基因座等位基因频率

等位基因	频率										
	中国汉族	中国汉族	中国汉族（北京）	中国汉族（北京）	中国汉族（太原）	中国汉族（河北）	中国汉族（太原）	中国回族（宁夏）	中国壮族（广西）	美国白人	美国黑人
7		0.005		0.0033		0.0028	0.0050	0.0049	0.0149		0.002
8	0.2682	0.2733	0.3604	0.3604	0.2673	0.2933	0.2736	0.2233	0.2612	0.143	0.033
9	0.1273	0.150	0.1599	0.1599	0.1238	0.1564	0.1965	0.1602	0.2164	0.052	0.019
10	0.1500	0.140	0.1066	0.1066	0.1485	0.1453	0.1567	0.1068	0.1791	0.052	0.026
11	0.2591	0.2467	0.1701	0.1670	0.2574	0.1872	0.1940	0.3107	0.1940	0.305	0.309
12	0.1682	0.140	0.1675	0.1670	0.1733	0.1704	0.1418	0.1262	0.0970	0.307	0.414
13	0.2270	0.035	0.0279	0.0281	0.0248	0.0363	0.0249	0.0485	0.0373	0.083	0.149
14	0.0045	0.010	0.0076	0.0077	0.0050	0.0084	0.0075	0.0194		0.057	0.049
H	0.793	0.837		0.7774	0.79		0.7910	0.7670	0.7761	0.71	0.69
DP	0.9259	0.928		0.9183	0.9149		0.9332	0.9192	0.9187		

1. 引物序列　5′-ACA GAA GTC TGG GAT GTG GA-3′

5′-GCC CAA AAA GAC AGA CAG AA-3′

2. PCR 反应　在 10μl 反应体系中，含有 10mmol/L Tris-HCl，pH 8.3，50mmol/L KCl，1.5mmol/L $MgCl_2$，200μmmol/L dNTPs，每种引物 1pmol，Taq DNA 聚合酶 0.3u，模板 DNA 30ng。

3. 热循环参数　94℃ 30s，55℃ 75s，72℃15s，27 个循环，最后 72℃保温 6min。

13.19 D13S325 基因座

D13S325 基因座定位于人类 13 号染色体（GDB：6867720，核心序列为 AGAT，已检出等位基因 7 个，片段大小为 211～235bp。

1. 引物序列　5′-TCC TTT AAG TGT TCT AGA GAG GAG G-3′

5′-TCT CTC TCA GAA GTT TGG AAG C-3′

2. PCR 反应　在 10μl 反应体系中，含有 10mmol/L Tris-HCl，pH 8.3，50mmol/L KCl，1.5mmol/L $MgCl_2$，200μmmol/L dNTPs，每种引物 1pmol，Taq DNA 聚合酶 0.3u，模板 DNA 30ng。

3. 热循环参数　94℃ 30s，55℃ 75s，72℃15s，27 个循环，最后 72℃保温 6min。

13.20 D16S539 基因座

D16S539 基因座定位于人类染色体 16q24qter，核心序列为 AGAT，已检出 7 个等位基因，核心序列重复次数为 5，8～15，等位基因片段长度大小范围为 264～304bp，频率见表 13-17。

表 13-17 D16S539 基因座等位基因频率

等位基因	频				率			
	中国汉族	中国汉族（北京）	中国汉族（河北）	中国汉族（太原）	中国壮族	中国壮族（广西）	美国白人	美国黑人
5								0.002
7					0.1485			
8	0.0139	0.0406	0.0112	0.0149	0.0634		0.026	0.023
9	0.3102	0.3046	0.2849	0.3020	0.0048	0.2132	0.107	0.205
10	0.1157	0.1218	0.1006	0.1139	0.0297	0.1397	0.079	0.093
11	0.2500	0.1980	0.2598	0.2624	0.094	0.2721	0.319	0.316
12	0.1991	0.2360	0.2123	0.1980	0.3118	0.2353	0.269	0.202
13	0.1065	0.0838	0.1285	0.1040	0.2922	0.1250	0.167	0.133
14	0.0046	0.0152	0.0028	0.0050	0.0347	0.0147	0.031	0.026
15	0.000				0.0198		0.002	
杂合度	0.765			0.78	0.839	0.8382	0.73	0.78
DP	0.9153			0.9263	0.9109	0.9131		

1. 引物序列　5′-GAT CCC AAG CTC TTC CTC TT-3′
5′-ACG TTT GTG TGT GCA TCT GT-3′

2. PCR 反应　在 10μl 反应体系中，含有 10mmol/L Tris-HCl，pH 8.3，50mmol/L KCl，1.5mmol/L $MgCl_2$，200μmmol/L dNTPs，每种引物 1pmol，Taq DNA 聚合酶 0.3u，模板 DNA 30ng。

3. 热循环参数　94℃ 30s，55℃75s，72℃15s，27 个循环，最后 72℃保温 6min。

13.21 D17S976 基因座

D17S976 基因座（GDB G00-250-233）为一复合 STR，含有 ATCA、ATCT 和 ACCT 重复单位。在最小和最大等位基因中发现有 ATC 三核苷酸，重复区域序列结构见表 13-18。

表 13-18 D17S976 基因座重复区域序列

等位基因	片段长度	序列
19.3	236bp	$(ATCA)_3-(ATCT)_2-ATC-(ATCT)_5-T-(ATCT)-ATCCT-(ATCT)_8$
21	241	$(ATCA)_1-(ATCT)_{10}-T-(ATCT)-ATCCT-(ATCT)_9$
21.3	244	$(ATCA)_3-(ATCT)_2-ATC-(ATCT)_5-T-(ATCT)-ATCCT-(ATCT)_{10}$
22	245	$(ATCA)_1-(ATCT)_{10}-T-(ATCT)-ATCCT-(ATCT)_{10}$
23	249	$(ATCA)_1-(ATCT)_{11}-T-(ATCT)-ATCCT-(ATCT)_{10}$
		$(ATCA)_1-(ATCT)_{10}-T-(ATCT)-ATCCT-(ATCT)_{11}$
24	253	$(ATCA)_1-(ATCT)_{11}-T-(ATCT)-ATCCT-(ATCT)_{11}$
25	257	$(ATCA)_1-(ATCT)_{12}-T-(ATCT)-ATCCT-(ATCT)_{11}$
26	261	$(ATCA)_1-(ATCT)_{13}-T-(ATCT)-ATCCT-(ATCT)_{11}$
27	265	$(ATCA)_1-(ATCT)_{12}-T-(ATCT)-ATCCT-(ATCT)_{13}$

续表

27.3	268	$(ATCA)_3-(ATCT)_2-ATC-(ATCT)_9-T-(ATCT)-ATCCT-(ATCT)_2-(ACCT)-(ATCT)_9$
28.3	272	$(ATCA)_3-(ATCT)_2-ATC-(ATCT)_{11}-T-(ATCT)-ATCCT-(ATCT)_2-(ACCT)-(ATCT)_8$
29.3	276	$(ATCA)_3-(ATCT)_2-ATC-(ATCT)_{10}-T-(ATCT)-ATCCT-(ATCT)_2-(ACCT)-(ATCT)_{10}$
30.3	280	$(ATCA)_3-(ATCT)_2-ATC-(ATCT)_{11}-T-(ATCT)-ATCCT-(ATCT)_2-(ACCT)-(ATCT)_{10}$
31.3	284	$(ATCA)_3-(ATCT)_2-ATC-(ATCT)_{12}-T-(ATCT)-ATCCT-(ATCT)_2-(ACCT)-(ATCT)_{10}$
32.3	288	$(ATCA)_3-(ATCT)_2-ATC-(ATCT)_{12}-T-(ATCT)-ATCCT-(ATCT)_2-(ACCT)-(ATCT)_{11}$

1. 引物序列　5′－ATA TGC CAC CAC TGG TT－3′

5′－TGG TAG CAT GCA TCT GTAGTC C－3′

2. PCR 反应　25μl 反应体系，含有 10mmol/L Tris－HCl（pH 9.0），50mmol/L KCl，1.5mmol/L $MgCl_2$，0.001% 白明胶，0.1% Triton－X 100，10μg BSA，200 μmol/L dNTPs，每种引物 0.4μmol/L，Taq DNA 聚合酶 0.5u，模板 DNA 5ng。

3. 热循环参数　94℃ 1min，66℃ 1min，72℃ 2min，6 个循环，93℃ 1min，65℃1min，72℃ 2min，6 个循环，93℃ 1min，63℃ 1min，72℃ 2min，18 个循环。

13.22 D18S51 基因座

D18S51 基因座定位于人 18 号染色体，核心重复序列为四核苷酸 AGAA，已检出 21 个等位基因，等位基因片段长度范围为 271～343bp，其中 18 个等位基因按一个四核苷酸重复单位递增，另三个等位基因（片段长度为 289，293，297bp）只增加 2bp，而不是增加 4bp，这三个等位基因的第一个四核苷酸 3′端旁侧序列有变化，有一个二核苷酸 AG 的插入（表 13－19），当这种插入发生在含有许多 AT 二核苷酸串联重复时就不知道这种插入发生的正确位置，这种变异的 3′旁侧可能是由于重复单位的第一个四核苷酸 A－G 的转化（AAAG－AGAG）造成的，这种突变的发生可能是通过滑动不匹配引起的。根据国际法医 DNA 委员会命名原则，以“.2”方式反映出这 2 个 bp，因此这三个等位基因命名为 13.2，14.2，15.2，这种情况与 FGA 基因座相似。表 13－20 为等位基因频率。

表 13－19　等位基因序列结构（M.D.Barber，1996）

等位基因	等位基因大小	重复序列	3′旁侧序列
9	271bp	(AGAA) 9	AAAG AGAGAG
10	275bp	(AGAA) 10	AAAG AGAGAG
11	279bp	(AGAA) 11	AAAG AGAGAG
12	283bp	(AGAA) 12	AAAG AGAGAG
13	287bp	(AGAA) 13	AAAG AGAGAG
13.2	289bp	(AGAA) 13	AGAGAGAGAGAG
14	291bp	(AGAA) 14	AAAG AGAGAG
14.2	293bp	(AGAA) 14	AGAGAGAGAGAG
15	295bp	(AGAA) 15	AAAG AGAGAG
15.2	297bp	(AGAA) 15	AGAGAGAGAGAG
16	299bp	(AGAA) 16	AAAG AGAGAG
17	303bp	(AGAA) 17	AAAG AGAGAG

续表

18	307bp	(AGAA) 18	AAAG AGAGAG
19	311bp	(AGAA) 19	AAAG AGAGAG
20	315bp	(AGAA) 20	AAAG AGAGAG
21	319bp	(AGAA) 21	AAAG AGAGAG
22	323bp	(AGAA) 22	AAAG AGAGAG
23	327bp	(AGAA) 23	AAAG AGAGAG
24	331bp	(AGAA) 24	AAAG AGAGAG
26	339bp	(AGAA) 26	AAAG AGAGAG
27	343bp	(AGAA) 27	AAAG AGAGAG

表 13-20 D18S51 基因座等位基因频率

等位基因	频率					
	中国汉族	中国汉族	中国汉族（北京）	中国汉族（重庆）	美国白人	美国黑人
10	0.0023				0.0050	0.0051
10.2						0.0051
11	0.0070	0.0017	0.0020	0.0200	0.0154	
12	0.0444	0.035	0.0383	0.0260	0.1425	0.0564
13	0.2126	0.2135	0.1800	0.1870	0.1500	0.0410
13.2			0.0083	0.0090	0.000	0.0051
14	0.2093	0.1817	0.2200	0.1760	0.1675	0.0667
14.2				0.0090		0.0051
15	0.1682	0.1967	0.1800	0.1900	0.1425	0.1769
15.2				0.0020		
16	0.1355	0.1183	0.1100	0.1500	0.1400	0.1744
17	0.0584	0.8330	0.0867	0.0820	0.1050	0.1641
18	0.0444	0.0467	0.0550	0.0460	0.0600	0.1103
19	0.0467	0.0400	0.0467	0.0580	0.0375	0.1026
20	0.1636	0.0217	0.0267	0.0170	0.0150	0.0487
21	0.0327	0.0333	0.0267	0.0190	0.0100	0.0103
22	0.0164	0.0200	0.0117	0.0100	0.0025	0.0103
23	0.0094	0.0017	0.0017	0.0100	0.0025	0.0026
24	0.0023	0.0050	0.0050	0.0030		
25				0.0030		
26		0.0017	0.0017			
杂合度	0.8551	0.8667	0.8603	0.8300		
DP	0.9603	0.9615	0.9818	0.9610		

1. 引物序列　5′-CAA ACC CGA CTA CCA GCA AC-3′

5′-GAG CCA TGT TCA TGC CAC TG-3′

2. PCR 反应　20μl 反应体系，含有 10mmol/L Tris-HCl（pH 9.0），50mmol/L KCl，1.5mmol/L $MgCl_2$，0.001% 白明胶，0.1% Triton-X 100，10μg BSA，200 μmol/L dNTPs，每种引物 8pmol，Taq DNA 聚合酶 0.8u，模板 DNA 5ng。

3. 热循环参数　94℃ 1min，59 ℃1min，72℃1min，28 个循环。

13.23 D18S535 基因座

D18S535 基因座（M.V.Lareu，1998）定位于人类染色体 18（GenBank：G00 – 208 – 299），核心序列为 GATA，在加利西亚人已检出 8 个等位基因，片段大小为 130 ~ 158bp。

1. 引物序列　5′ – TCA TGT GAC AAA AGC CA AC – 3′（5′ – 末端用荧光标记）

　　5′ – AGA CAG AAA TAT AGA TGA GAA TGC A – 3′

2. PCR 反应体系　50μl 反应体系，含有 10mmol/L Tris – HCl（pH 8.3），50mmol/L KCl，1.5mmol/L $MgCl_2$，0.001% 白明胶，0.1% Triton – X 100，200μmol/L dNTPs，每种引物 0.25μmol/L，Taq DNA 聚合酶 1.25u，模板 DNA 5ng。

3. 热循环参数　94℃变性 45s，60 ℃1min，72℃1min，30 个循环。

13.24 D18S849 基因座

D18S849 基因座定位于人类染色体 18q12 – q21，已检出 9 个等位基因，等位基因片段长度大小范围为 93 ~ 133bp，杂合度为 0.758，非父排除率为 0.539。

1. 引物序列　5′ – AGG TCC CAG CCA TAT AGA CG – 3′

　　5′ – GCC CTA GAA CCA GGG ATT AA – 3′

2. PCR 反应体系　在 10μl 反应体系中，含有 10mmol/L Tris – HCl（pH 8.3），50mmol/L KCl，1.5mmol/L $MgCl_2$，200μmmol/L dNTPs，每种引物 1pmol，Taq DNA 聚合酶 0.3u，模板 DNA 30ng。

3. 热循环参数　94℃ 30s，55℃ 75s，72℃15s，27 个循环，最后 72℃保温 6min。

13.25 D19S253 基因座

D19S253 基因座定位于 19 号人染色体，核心序列为 GATA，已检出 10 个等位基因，等位基因片段大小为 209 ~ 245bp。表 13 – 21 为 D19S253 基因座等位基因频率。

表 13 – 21　D19S253 基因座等位基因频率

等位基因	频率		
	中国汉族（成都）	中国汉族（广州）	中国壮族
7	0.1714	0	0.1485
8	0.0714	0.124	0.0643
9	0.0000	0.019	0.0048
10	0.0333	0.014	0.0297
11	0.0952	0.005	0.094
12	0.3000	0.152	0.3118
13	0.1810	0.286	0.2922
14	0.1334	0.262	0.0347
15	0.0143	0.109	0.0198
16		0.029	
H	0.76	0.8353	0.8390
DP	0.82	0.9211	0.9109
PE		0.7499	0.8367

1. 引物序列 5′ - ATA GAC AGA CAC ACG GAC TG - 3′

5′ - GGG AGT GGA CAT TAC CCC T - 3′

2. PCR 反应体系 30μl 反应体系，含 50mmol/L KCl，10mmol/L Tris - HCl（pH 8.3），1.5mmol/L $MgCl_2$，dNTPs 20μmol/L，0.5u Taq DNA 聚合酶，模板 DNA 2μl（5～10ng），每种引物 0.1μmol/L。

3. 热循环参数 95℃ 2min，然后 94℃ 60s，60℃ 60s，72℃ 60s，循环 30 次。最后 72℃ 10min。

13.26 D19S400 基因座

D19S400 基因座定位于人类第 19 号染色体，核心序列为（GATA）n，已检出 11 个等位基因，等位基因片段大小为 185～225bp 之间，等位基因群体分布频率见表 13 - 22。

表 13 - 22 D19S400 基因座等位基因频率

等位基因	频率								
	中国汉族（成都）	中国汉族（广州）	中国汉族（吉林）	中国藏族（拉萨）	中国白族（大理）	中国壮族（南宁）	中国蒙古族	日本人	德国人
7		0.0051						0.0042	
8		0.005						0.0085	0.0104
9	0.228	0.1786	0.2071	0.1966	0.1477	0.1932	0.1044	0.2034	0.1042
10	0.0227	0.0357	0.0354	0.0618	0.0511	0.0739	0.0714	0.0636	0.0469
11	0.0519	0.1071	0.0707	0.0562	0.0739	0.0682	0.0659	0.0890	0.0729
12	0.3442	0.3061	0.3030	0.3596	0.3693	0.2273	0.3242	0.3136	0.1667
13	0.1494	0.1582	0.1869	0.173	0.1477	0.1307	0.1923	0.1059	0.1510
14	0.0942	0.0765	0.0859	0.0393	0.0398	0.0625	0.1209	0.0466	0.1667
15	0.0682	0.0714	0.0707	0.1124	0.0739	0.1307	0.0879	0.1059	0.1719
16	0.0455	0.0459	0.0404	0.0112	0.0795	0.0852	0.0275	0.0169	0.0938
17	0.0032	0.0102		0.0056	0.0170	.0284	0.0055		0.0156
H	0.75	0.77	0.78	0.89	0.75	0.81	0.81	0.84	0.88
DP	0.9301	0.9482	0.9394	0.9382	0.9625	0.9439	0.9480	0.9664	0.9664
CE	0.6057	0.6586	0.6320	0.5974	0.6268	0.71	0.6447	0.6582	0.7257

1. 引物序列 5′ - CGG TAT GTC TTT ATC AGC AG - 3′

5′ - ATG ACA GCT CTA GGA AGG C - 3′

2. PCR 反应体系 2～40ng 模板 DNA，1.25u Taq DNA 聚合酶，dNTPs 20μmol/L，0.25μmol/L 引物，1×PCR 缓冲液。

3. 热循环参数 94℃变性 5min 后，在 72℃条件下加引物，dNTPs 和 Taq DNA 聚合酶。循环条件：94℃ 1min，58℃ 1min，72℃ 2min，35 个循环后在 72℃保持 5min。

13.27 D20S161 基因座

D20S161 基因座定位于人 20 号染色体，重复序列为（TAGA，TAGG）n，序列为 CCCTTTC AACTTGCATG CTTTAATGAT AAAGAAATAC GGGTGTAGAA AAGTTTAAGTG ATTTGTC-

CAC CATGACA (TAGA)$_m$ (TAGG)$_n$ (TAGA)$_p$ TGGAATA AATGATTATA GGGCAAT-GAACCAGTTGGAAGGA，共有 8 个等位基因。群体分布频率见表 13 - 23。

表 13 - 23　D20S161 基因座等位基因频率

等位基因	频率							
	中国汉族（成都）	中国汉族（广州）	中国汉族（吉林）	中国藏族（拉萨）	中国白族（大理）	中国蒙古族	南宁壮族	德国人
14								
16	0.0491	0.0368	0.0657	0.0220	0.309	0.0459	0.0500	0.0200
17	0.2254	0.2947	0.1970	0.1813	0.2216	0.1429	0.200	0.1750
18	0.3237	0.3421	0.4040	0.4835	0.3763	0.4184	0.3550	0.5100
19	0.1532	0.1421	0.0909	0.1099	0.1598	0.1326	0.1850	0.2050
20	0.1301	0.1369	0.1263	0.1429	0.1185	0.1786	0.1300	0.0600
21	0.0983	0.0421	0.1010	0.0439	0.0928	0.0612	0.0550	0.0250
22	0.0202	0.0053	0.0151	0.0165		0.0204	0.0250	0.0050
H	0.7514	0.7368	0.7273	0.7253	0.8041	0.6531	0.7600	0.7300
Dp	0.9276	0.9010	0.9124	0.8741	0.9083	0.9062	0.9192	0.8427
CE	0.5973	0.5337	0.5596	0.4817	0.5489	0.5450	0.5758	0.4287

1. 引物序列　5′ - CCC TTT CAA CTT GCA - 3′
　　5′ - TCC TTC CAA CTG GTT CAT - 3′

2. PCR 反应条件　25μl 反应体系，含 50mmol/L KCl，10mmol/L Tris - HCl（pH 8.3），1.5mmol/L $MgCl_2$，150μmol/ L dNTPs，1u Taq DNA 聚合酶，5 ~ 10ng 模板 DNA，0.25μmol/L 引物。

3. 热循环参数　94℃ 60s，60℃60s，72℃ 60s，循环 32 次。

13.28 D20S470 基因座

D20S470 基因座（GDB：686139）定位于人类 20 号染色体，重复序列为 TTCC，已检出 12 个等位基因，片段大小为 277 ~ 321bp（Dong Hoon Lee，1998）。

1. 引物序列　5′ - CCT TGG GGG ATA TAG CCT AA - 3′
　　5′ - TGA GTG ACA GAG TGA TAC CAT G - 3′

2. PCR 反应体系　在 10μl 反应体系中，含有 10mmol/L Tris - HCl（pH 8.3），50mmol/L KCl，1.5mmol/L $MgCl_2$，200μmmol/L dNTPs，1pmo 引物 1，0.3u Taq DNA 聚合酶，10ng 模板 DNA。

3. 热循环参数　94℃ 30s，55℃ 75s，72℃延伸 15s，27 个循环，最后 72℃保温 6min。

13.29 D21S11 基因座

D21S11 基因座定位于人类第 21 号染色体，重复序列为 TCTA/TCTG，已检出 22 个等位基因，等位基因片段大小为 172 ~ 264bp。Moller（1994）用 Sharma Litt（1992）引物扩增，并对扩增进行测序分析发现，在重复序列内有 3 个可变区（表 13 - 24）和一个恒定区，在等位基因 29 后的等位基因存在一不完全重复子，可能是由于在 3′端最后二个重复

子中 TA 插入或 TC 缺失造成的。在可变Ⅰ区由（TCTA）4～6 组成，Ⅱ区为（TCTG）5～6 组成，在Ⅱ区和Ⅲ区之间有二、三和六核苷酸组成，Ⅲ区重复数目变化最大，目前凝胶电泳方法还区分不开 29 和 31 结构变异体。群体等位基因分布频率见表 13－25。

表 13－24 D21S11 基因座重复子结构、片段长度和等位基因命名

等位基因命名	片段长度	序列结构 Ⅰ	Ⅱ	Ⅲ
26	209	$(TCTA)_4$	$(TCTG)_6$ * * * * * * * * *	$(TCTA)_8$
27	213	$(TCTA)_4$	$(TCTG)_6$ * * * * * * * * *	$(TCTA)_9$
28	217	$(TCTA)_4$	$(TCTG)_6$ * * * * * * * * *	$(TCTA)_{10}$
29	221	$(TCTA)_4$	$(TCTG)_6$ * * * * * * * * *	$(TCTA)_{11}$
29'	221	$(TCTA)_6$	$(TCTG)_5$ * * * * * * * * *	$(TCTA)_{10}$
30	225	$(TCTA)_6$	$(TCTG)_5$ * * * * * * * * *	$(TCTA)_{11}$
31'	229	$(TCTA)_6$	$(TCTG)_5$ * * * * * * * * *	$(TCTA)_{12}$
31	229	$(TCTA)_5$	$(TCTG)_6$ * * * * * * * * *	$(TCTA)_{12}$
32	233	$(TCTA)_6$	$(TCTG)_5$ * * * * * * * * *	$(TCTA)_{13}$
30.2	227	$(TCTA)_5$	$(TCTC)_6$ * * * * * * * * *	$(TCTA)_{10}$TATCTA
31.2	231	$(TCTA)_5$	$(TCTG)_6$ * * * * * * * * *	$(TCTA)_{11}$TATCTA
32.2	235	$(TCTA)_5$	$(TCTG)_6$ * * * * * * * * *	$(TCTA)_{12}$TATCTA
33.2	239	$(TCTA)_5$	$(TCTG)_6$ * * * * * * * * *	$(TCTA)_{13}$TATCTA
34.2	243	$(TCTA)_5$	$(TCTC)_6$ * * * * * * * * *	$(TCTA)_{14}$ TATCTA

"* * * * * * * * *"部分为恒定区（43bp），序列为（TCTA）3TA（TCTA）3TCA（TCTA）2TCCATA

表 13－25 D21S11 基因座等位基因频率

等位基因	频率 中国汉族	中国汉族（北京）	中国汉族（广州）	中国汉族（重庆）	中国壮族	美国白人	美国黑人
<26						0.0050	
26							0.0026
27	0.0017	0.0017		0.0070		0.0375	0.0590
28	0.0350	0.0533	0.0198	0.0540		0.1625	0.2180
28.2	0.0050	0.0183	0.0049	0.0100			
29	0.2333	0.2833	0.1040	0.2450	0.034	0.2075	0.2000
29.2		0.0083	0.0099	0.0100	0.0291	0.0050	
29.3							0.0026
30	0.0017	0.2733	0.2525	0.2680	0.2476	0.2625	0.1615
30.2	0.2823	0.0117		0.0170	0.0243	0.0250	0.0256
31	0.0067	0.1017	0.2129	0.1090	0.2476	0.0550	0.0897
31.2	0.1033	0.0583	0.0099	0.0770	0.0534	0.1050	0.0590
31.3	0.0750						
32	0.0300	0.0383	0.1188	0.0200	0.0583	0.0125	0.0077
32.2	0.1383	0.1000	0.0792	0.1190	0.0291	0.0725	0.0692
33	0.0017	0.0033	0.0445	0.0070	0.0680	0.0025	0.0051

续表

33.1			0.000				0.0026
33.2	0.0482	0.0450	0.1045	0.0540*	0.0340	0.0400	0.0410
34					0.1359		0.0026
34.2	0.0083	0.0033	0.0347			0.0075	
35					0.0291		0.0359
35.2			0.0049		0.0049		
36			0.0049		0.0154		
38							0.0026
杂合度	0.8433	0.8157	0.8974	0.8090	0.8556		
DP	0.9416	0.9818	0.8429	0.9460	0.8680		
PIC	0.8383	0.7922	0.81	0.8844			
PE	0.7179	0.6454			0.8784		

* 为等位基因33.2后的等位基因总频率

1. 引物序列　5′－ATA TGT GAG TCA ATT CCC CAA G－3′

5′－TGT ATT AGT CAA TGT TCT CCAG－3′

2. PCR反应体系　50μl反应体系，含50mmol/L KCl，10mmol/L Tris－HCl（pH 8.3），1.5mmol/L $MgCl_2$，0.01mg/ml BSA，dNTPs 160μmol/L，1u Taq DNA聚合酶，5～10ng模板DNA，0.25μmol/L引物。

3. 热循环参数　94℃变性2min后，94℃ 45s，60℃ 30s，72℃ 45s，循环30次，72℃保温10min。

或　Sharma Litt（1992）设计的引物位于靶序列5′旁侧31bp，3′旁侧63bp：

引物序列　5′－GTG AGT CAA TTC CCC AAG－3′

5′－GTT GTA TTA GTC AAT GTT CTCC－3′

13.30 D22S444基因座

D22S444基因座，定位于22号染色体。核心序列ATCC。中国南方群体个人识别率0.621，PIC 0.36，非父排除率0.120，杂合度0.468；北方群体个人识别率0.444，PIC 0.24，非父排除率0.078，杂合度0.280。

1. 引物序列　5′－TTT CAA CTA AGC CTT AAA AAT GC－3′

5′－TGT TTG GCT TGA AGA AGG AG－3′

2. PCR反应体系　体积37.5μl，3ng DNA模板，1mmol/dNTPs L、1.5u Taq DNA聚合酶，1.5mmol/L $MgCl_2$，BSA 3.75μl（1.6μg/ml），50nmol/L引物。

3. 热循环参数　94℃ 50s，55℃ 50s，72℃ 30s，38个循环。

13.31 D22S445基因座

D22S445基因座，核心序列ATCC。中国南方群体个人识别率0.850，PIC 0.68，非父排除率0.490，杂合度0.682；北方群体个人识别率0.822，PIC 0.59，非父排除率0.402，杂合度0.645。

1. 引物序列 5′ - GTC CAT CCG TTT GTT TGT TC - 3′

5′ - TGG ATG GAG AGA AGG AAT GA - 3′

2. PCR反应体系、热循环参数及电泳分型同D22S444。

13.32 D22S533 基因座

D22S533基因座，核心序列ATAG。中国南方群体个人识别率0.909，PIC 0.74，非父排除率0.401，杂合度0.78；北方群体个人识别率0.889，PIC 0.73，非父排除率0.647，杂合度0.825。

1. 引物序列 5′ - CCA CCT TCA ACC CTC CAA T - 3′

5′ - TTC TTA GTG GGG AAG GCA TC - 3′

2. PCR反应体系、热循环参数及电泳分型同D22S444。

13.33 D22S683 基因座

D22S683基因座，核心序列TATC。中国南方群体个人识别率0.960，PIC 0.86，非父排除率0.472，杂合度0.727；北方群体个人识别率0.957，PIC 0.84，非父排除率0.617，杂合度0.810。

1. 引物序列 5′ - AAC AAA ACA AAA CAA ACA - 3′

5′ - GGT GGA AAT GCC TCA TGT AT - 3′

2. PCR反应体系、热循环参数及电泳分型同D22S444。

13.34 D22S685 基因座

D22S685基因座，核心序列GATA。中国南方群体个人识别率0.909，PIC 0.75，非父排除率0.633，杂合度0.818；北方群体个人识别率0.923，PIC 0.77，非父排除率0.617，杂合度0.810。

1. 引物序列 5′ - TTC TTA GTG GGG AAG GGA TC - 3′

5′ - TGA TTT GAT GTT TTT GAT AGA CA - 3′

2. PCR反应体系、热循环参数及电泳分型同D22S444。

13.35 D22S686 基因座

D22S686基因座，核心序列CCTT。中国南方群体个人识别率0.860，PIC 0.68，非父排除率0.435，杂合度0.705；北方群体个人识别率0.860，PIC 0.70，非父排除率0.587，杂合度0.794。

1. 引物序列 5′ - TTG ATT ACA GAG TGG CTC TGG - 3′

5′ - TAA GCC CTG TTA GCA CCA C - 3′

2. PCR反应体系、热循环参数及电泳分型同D22S444。

13.36 ACTBP2 基因座

ACTBP2基因座即SE33基因座，是人类的β-肌动蛋白基因座有关的一个假基因（hu-

man beta - actin related psendogene）（GenBank accession No.v00481）。核心序列为 AAAG，已检出 26 个等位基因，基因座杂合度 0.93，识别率 0.99。位于 223～259bp 范围的等位基因含有有规律的 4 bp 的 AAAG 重复结构，大于等位基因 19 的等位基因序列结构有差异，在大片段中可能由于 AA 的插入含有一个 AAAAAG，这个六聚体的位置可变化，而且在一些等位基因内，六聚体内发生 A→G 置换，变成 AGAAAG 序列（如图 13 - 1，2 中等位基因 26 和 31）；在 311～323bp 范围的等位基因中，在重复区域内有 2 个不同的六核苷酸单位，致使多态性增加。最小等位基因 7（202bp）只有 7 个（AAAG）重复，在 3′旁侧区接近重复区有 1bp G 缺失。等位基因频率分布见表 13 - 26。

表 13 - 26　ACTBP2 基因座等位基因频率

等位基因	频率		
	中国汉族（成都）	美国人	德国人
11			0.019
12			0.019
13		0.01	0.033
14		0.05	0.022
15		0.05	0.050
16	0.0034	0.04	0.072
17	0.0068	0.02	0.053
18	0.0068	0.06	0.058
19	0.0102	0.14	0.025
20	0.0408	0.04	0.042
21	0.0782	0.01	0.008
22	0.0952	0.03	0.025
23	0.0714	0.04	0.017
24	0.0510	0.03	0.050
25	0.0816	0.03	0.031
26	0.0578	0.04	0.036
27	0.0408	0.03	0.033
28	0.0510	0.09	0.083
29	0.0544	0.09	0.058
30	0.0578	0.10	0.092
31	0.1054	0.05	0.044
32	0.0578	0.04	0.047
33		0.01	0.022
34	0.0816		0.044
35			0.003
36	0.0442		0.011
38	0.0034		
杂合度			0.93
DP			0.99

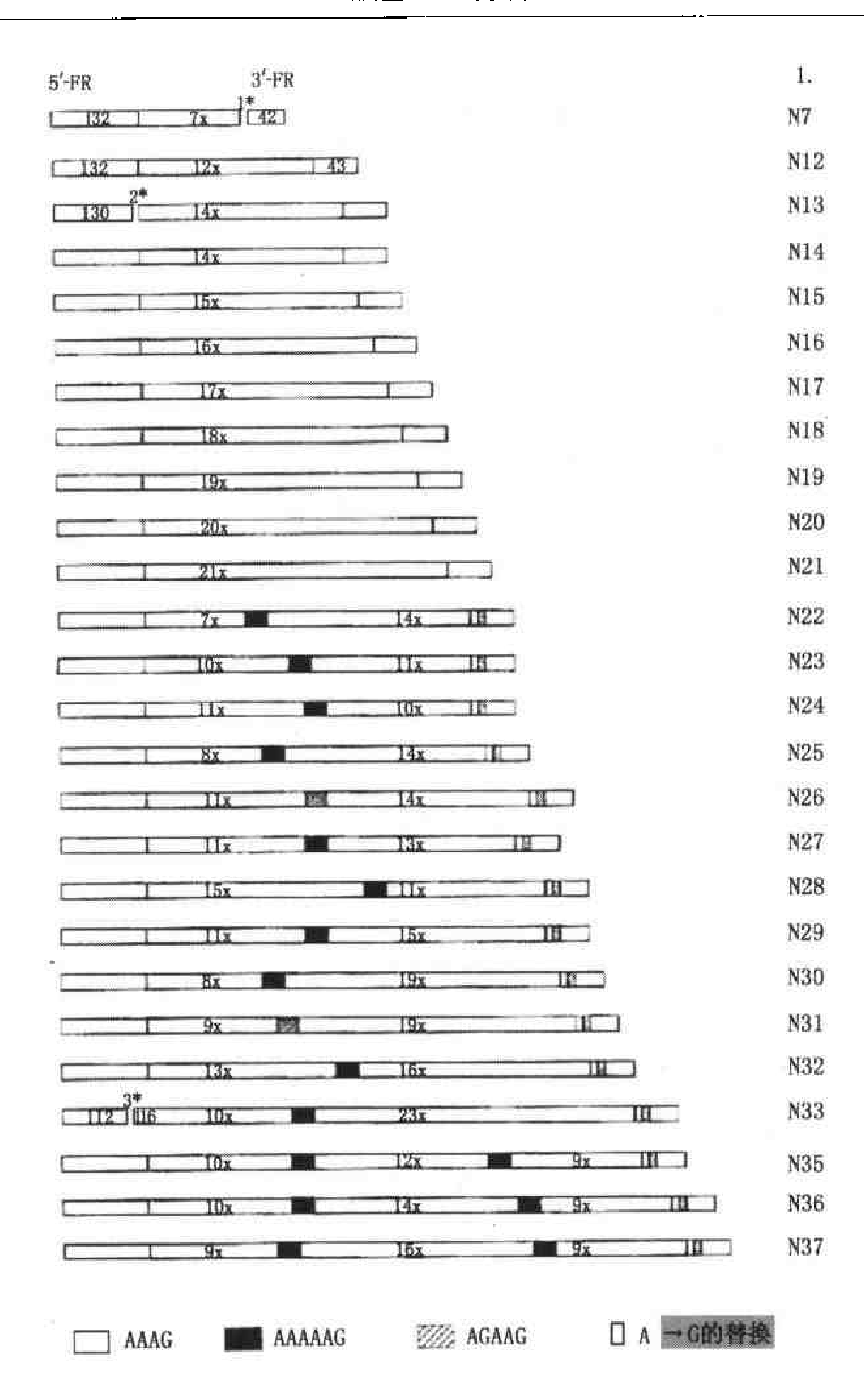


图 13－1　ACTP2 ladder 的 26 个等位基因序列结构

1＊ 在重复区附近有 1 个 bp 的缺失，2＊在 131－132 位置存在 AG 缺失，
3＊在 113－116 位置有 AAAG 的缺失

1. 引物序列　5'－AAT CTG GGC GAC AAG AGT GA－3'
　　　　　　5'－ACA TCT CCC CTA CCG CTA TA－3'

2. PCR 反应体系　25μl 反应体系，含有 1ng 模板 DNA，1u Taq DNA 聚合酶，0.3μmol/L 引物，100μmol/L dNTPs，2.5μl 10×PCR 缓冲液。

3. 热循环参数　93℃1min，61℃1min，72℃1.5min，28 个循环。

13.37 HUMCD4 基因座

HUMCD4 基因座为识别/表面抗原（recognition/surface antigen，CD4），定位于人类染色体 12p，重复序列为 AAAAG，已检出 8 个等位基因，杂合度为 0.7071。引物对 1 扩增等位基因片段大小为 140～170bp，引物对 2 的扩增片段大小为 88～128 bp。群体分布频率见表 13－27。

表 13－27　HUMCD4 基因座等位基因频率

等位基因	频率			
	中国汉族（台湾）	澳地利人	美国白人	美国黑人
5	0.00	0.337	0.365	0.286
6	0.005	0.331	0.295	0.214
7	0.54	0.003	0.005	0.107
8	0.015	0.003	0.005	0.036
9		0.010	0.300	0.036
10		0.278	0.025	0.250
11	0.01	0.025	0.005	0.036
12	0.42	0.013		0.036
13	0.01			
杂合度	0.54	0.7071		
DP	0.678	0.845		
PIC		0.638		

1. 引物序列　5′－TTA CGC GTT TGG AGT CGC AAG CTG AAC TAG CG－3′
　　5′－CCA GGA AGT TGA GGC TGC ACT GAA－3′
　或　5′－TTG GAG TCG CAA GCT GAA CTA GC－3′
　　5′－GCC TGA GTG ACA GAG TGA GAA CC－3′

2. PCR 反应体系　25μl 反应体系，含有 50mmol/L KCl，10mmol/L Tris－HCl（pH 9.0），1.5mmol/L $MgCl_2$，0.001% 白明胶，200μmol/LdNTPs，引物 1.5pmol，1u Taq DNA 聚合酶，5～15ng DNA 模板。

3. 热循环参数　94℃1min，60℃45s，72℃1min 30s，30 个循环。

13.38 HUMCSF1PO 基因座

HUMCSF1PO 基因座是人 c－fms 原癌基因（c－fms proto－oncogene），也是 CSF－1 受体（CSF－1 receptor）、又叫 CSF1R，定位于人类染色体 5q33.5－q34，重复单位为 AGAT，已检出等位基因 6～14 等 10 个等位基因，片段长度为 295～327bp，等位基因群体分布频率见表 13－28。

表 13－28 HUMCSF1PO 基因座等位基因频率

等位基因	频率							
	中国汉族（北京）	中国汉族（上海）	中国汉族（武汉）	中国汉族（青岛）	中国汉族（台湾）	中国壮族（广西）	中国苗族（云南）	西班牙人
7			0.005		0.013	0.0149		
8	0.019		0.002	0.018	0.004	0.0050	0.0036	0.003
9	0.0833	0.0507	0.048	0.0721	0.056	0.0693	0.0109	0.011
10	0.25	0.2765	0.205	0.2433	0.272	0.1733	0.2065	0.267
11	0.3333	0.2350	0.234	0.2703	0.198	0.2921	0.2536	0.275
12	0.025	0.2949	0.402	0.3198	0.409	0.2673	0.3841	0.374
13	0.0571	0.1152	0.093	0.0676	0.047	0.1238	0.1087	0.067
14	0.0071	0.0092	0.011	0.009		0.0446	0.0290	0.003
15		0.0138				0.0099	0.0036	
16		0.0047						
杂合度	0.7532	0.798	0.7316	0.7748	0.771	0.7030	0.6812	0.711
DP	0.8921	0.882	0.8901	0.8998		0.9553	0.8858	

1. 引物序列 5′－AAC CCT GAG TCT GCC CAA GGA CTA GC－3′

5′－TTC CAC ACA CCA CTG GCC ATC TTC－3′

或 Hammond 设计的引物，扩增产物片段大小为 319－347bp。

5′－GTT GCT AAC CAC CCT GTG TCT－3′

5′－TTC CTG TGT CAG ACC CTG TTC－3′

2. PCR 反应体系 25μl 反应体系，50mmol/L KCl，10mmol/L Tris－HCl（pH 9.0），1.5mmol/L $MgCl_2$，0.001%白明胶，200μmol/L dNTPs，0.5μmol/L 引物，0.25u Taq DNA 聚合酶。

3. 热循环参数 96℃变性 2min，94℃ 1min，58℃ 1min，70℃ 1.5min，10 个循环后，90℃ 1min，58℃1min，70℃1.5min，20 个循环，72℃5min。

13.39 HUMCYAR04（CYP19）基因座

HUMCYAR04（CYP19）基因座是位于芳香化酶细胞色素 P－450 基因（aromatase cytochrome P－450）内含子，定位于人染色体 15q21.1。重复单位核心序列为 AAAT。已检出 7 个等位基因（5～12）。该基因座片段长度范围 173～201bp，群体等位基因分布频率见表 13－29。

表 13－29 HUMCYAR04 基因座等位基因频率

等位基因	频率		
	中国汉族（云南）	中国汉族（台湾）	日本人
5	0.1073		
6	0.3272	0.400	
7	0.1597	0.0170	0.5875
8	0	0.000	0.0025
9	0.0209	0.000	0.000

续表

10	0.3351	0.010	0.0025
11	0.0471	0.385	0.325
12	0.0026	0.035	0.080
13	0.000	0.000	0.0025
杂合度	0.6859	0.59	0.565
DP	0.8879	0.82	0.701

1. 引物序列　5′ – GGT AAG CAG GTA CTT AGT TAG CTA C – 3′

5′ – GTT ACA GTG AGC CAA GCT CGT GAC – 3′

2. PCR 反应体系　25μl 反应体系，含有 50mmol/L KCl，10mmol/L Tris – HCl（pH 8.3），200μmol/L dNTPs，1.5mmol/L $MgCl_2$，0.75u *Taq* DNA 聚合酶，10ng 模板 DNA，0.2pmol/L 引物。

3. 热循环参数　95℃30s，65℃30s，72℃40s，30 个循环，72℃5min。

13.40 DHFRP2 基因座

DHFRP2 基因座（human dihydrofolate reductase psi – 2 pseudogene ，DHFRP2）定位于人类第 6 号染色体，重复单位为 AAAC，已检出 6 个等位基因，等位基因片段大小为 157 ~ 173bp。群体等位基因分布频率见表 13 – 30。该基因座的基本序列为：

GTAAGACTTTTGGAGCCATTTGATTTTACTTAACCATTTTCATGTA

TCCTTTAATAAAAACAACTCTATCATAATAAAAAGTCACACTTGTTCATAGA（AAAC）n

AAAAAAACCTTCCTGCCCAGCCCATCTCATTCTCCCTGAA

表 13 – 30　DHFRP2 基因座等位基因频率

等位基因	频率				
	中国汉族（成都）	中国汉族	美国白人	澳大利亚人	日本人
6	0.0865	0.100	0.04	0.150	0.050
7	0.4359	0.350	0.44	0.350	0.450
8	0.266	0.250	0.15	0.550	0.300
9	0.0449	0.150	0.25	0.050	0.050
9.5	0.0256		0.00	0.050	0.050
10	0.141	0.150	0.02		0.100
11				0.000	
杂合度	0.673	0.76	0.70	0.630	0.690
DP	0.87	0.905		0.819	0.857
CE	0.455	0.540		0.397	0.454

1. 引物序列　5′ – GTA AGA CTT TTG GAG CCA TT – 3′

5′ – TTC AGG GAG AAT GAG ATG GG – 3′

2. PCR 反应体系　50μl 反应体系，含有 50mmol/L KCl，10ng 模板 DNA，10mmol/L Tris – HCl（pH8.3），1.5mmol/L $MgCl_2$，0.001% 白明胶，200μmol dNTPs，0.1μmol/L 引物，0.5u Taq DNA 聚合酶。

3. 热循环参数　95℃30s，62℃30s，72℃40s，30 个循环，72℃5min。

13.41 GABRB15 基因座

GABRB15 基因座在中国汉族中已检出 6 个等位基因（表 13 – 31），片段大小为 150 ~ 170bp。

表 13 – 31　GABRB15 基因座等位基因频率

等位基因*	频　率
n1	0.0475
n2	0.4100
n3	0.3475
n4	0.1150
N5	0.0700
N6	0.0100
杂合度	0.6907
DP	0.8273
PIC	0.6380
PE	0.4436

* 按等位基因片段大小顺序命名

1. 引物序列　5′ – CTA GAA AGAC TAG CAA GGT GGA T – 3′
　　　　　　5′ – GCT CAT TAA ACA CTG TGT TCC T – 3′

2. PCR 反应体系　20μl 反应体系，含有 50mmol/L KCl，10mmol/L Tris – HCl（pH8.3），1.5mmol/L $MgCl_2$，0.001% 白明胶，0.2mmol/L dNTPs，0.04μmol/L 引物，8μl Chelex – 100 提取 DNA 溶液，0.5u Taq DNA 聚合酶。

3. 热循环参数　94℃1min，60℃1min，70℃1.5min，10 个循环；90℃1min，60℃1min，70℃1.5min，20 个循环，最后 60℃10min。

13.4 HUMFABP 基因座

HUMFABP 基因座是人肠脂肪酸结合蛋白（intestinal fatty acid – binding protein）基因内含子 STR 系列，定位于人染色体 4q28 – q31（GenBank Accession no. M18079），核心序列为 ATT。已检出 9 个等位基因，等位基因片段大小为 214 ~ 238bp，群体分布频率见表 13 – 32。

表 13－32　HUMFABP 基因座等位基因频率

等位基因	频	率	
	中国汉族（台湾）	德国人	美国人
9	0.07		
10	0.50	0.471	0.497
11	0.045	0.184	0.174
12	0.035		0.034
13	0.30	0.38	0.248 ·
14	0.045		0.044
15	0.005	0.009	
杂合度	0.79		
DP	0.825		

1．引物序列　5′－GTA GTA TCA GTT TCA TAG GGT CAC C －3′

5′－CAG TTC GTT TCC ATT GTC TCT CCG－ 3′

2．PCR 反应体系　25μl 反应体系，含有，50mmol/L KCl，10mmol/L Tris－HCl（pH 8.3），1.5mmol/L $MgCl_2$，0.001%白明胶，187.5μmol/L dNTPs，1.4μmol/L 引物，0.5u Taq DNA 聚合酶，1ng 模板 DNA。

3．热循环参数　95℃45s，65℃30s，72℃30s，25 个循环后，72℃5min。

13.43 HUMFES/FPS 基因座

HUMFES/FPS 基因座是 c－fes/fps 原癌基因（c－fes/fps proto－oncogene），定位于染色体 15q25－qter，（GenBank accession No.X06292，）含有一个简单的四核苷酸重复序列 ATTT，已检出 10 个等位基因，重复数为 8～13，片段长度为 222～250bp，在一些等位基因 10 和 11 的 5′－侧翼序列（155bp）的第 34 个核苷酸位置 A 变成 C（图 13－2），24bp 的 3′－旁侧则没有变化，发生这种变化的机率在等位基因 10 为 3/8，等位基因 11 则为 1/2。FES/FPS 基因座频率分布见表 13－33。

等位基因命名 (1)	片段长度 (bo) (2)	序列结构 5′-FR* 3′-FR**	测序等位基因数
8	211	A (ATTT) 9	2
9	215	A (ATTT) 9	1
10a	219	C (ATTT) 10	3
10	219	A (ATTT) 10	5
11a	223	C (ATTT) 11	5
11	223	A (ATTT) 11	6
12	227	A (ATTT) 12	2
13	231	A (ATTT) 13	1

图 13－2　FES/FPS 基因座重复序列结构

表 13－33　FES/FPS 基因座等位基因频率

	频			率				
等位基因	中国汉族（北京）	中国汉族（吉林）	中国汉族（广州）	中国汉族（太原）	中国壮族（云南）	中国景颇族（云南）	美国白人	美国黑人
6	0	0.005						
7	0.0025						0.014	0.003
8	0.0075			0.0092		0.0094		0.073
9	0.0050	0.020	0.0152	0.0275	0.0048	0.0236	0.307	00.046
10	0.6040	0.0202	0.0455	0.0550	0.0667	0.0519	0.39	024
11	0.5125	0.5101	0.4735	0.4404	0.3810	0.4434	0.22	0.357
12	0.2625	0.2424	0.3258	0.2936	0.3333	0.2925	0.063	0.23
13	0.1625	0.1869	0.1326	0.1606	0.2048	0.1651	0.005	0.048
14	0.0075	0.0101	0.0007	0.0138	0.0095	0.0142		
15		0.0050						
H		0.6484	0.5909	0.8921	0.7006	0.6887		
DP		0.8205	0.8165	0.8644	0.8566			
PM			0.1835	0.1356				
PIC			0.5889					

1. 引物序列　5′－GCT TGT TAA TTAC ATC TAC GGA AGGC－3′

　　　　　　5′－GTA GTC CCA GCT ACT TGG CTA CTC－3′

　　　　或　5′－GGG ATT TCC CTA TGG ATT GG－3′

　　　　　　5′－GCG AAA GAA TGA GAC TAC AT－3′

2.PCR 反应体系　25μl 反应体系，含有 1～5ng 模板 DNA，50mmol/L KCl，10mmol/L Tris－HCl（pH 8.3），1.5mmol/L $MgCl_2$，0.001%白明胶，200μmol/L dNTPs，1.4μmol/L 引物，1.0u Taq DNA 聚合酶。

3. 热循环参数　95℃1min，55℃1min，72℃1min，30 个循环后，72℃10min。

13.44 Hum FGA 基因座

Hum FGA 基因座即 HUMFIBRA，是人 α 血纤维蛋白基因（human alpha fibrinogen）第三内含子内的 STR 序列，定位于染色体 4q28，核心序列为 TCTT 或表达为 CTTT，已检出 22 个等位基因，片段长度范围为 168～294bp。在 168～224bp 大小范围内的等位基因，大多数按 4bp 有规律的增加，最小的二个等位基因（168 和 176bp）之间相差 8bp，可以推断之间应有一个大小为 172bp 的等位基因存在，但至今尚没有观察到此等位基因；在 168～220bp 大小范围的等位基因表现为单一的四核苷酸（TCTT）重复单位，其中有 5 个等位基因（182，190，198，202，206bp）在大小上与正常的 4bp 增加有差异，相差 2bp，在这 5 个等位基因中，每一个等位基因的第一个 TCTT 重复单位前没有 TC 二核苷酸存在，这可能是 TC 二核苷酸的缺失或 TT 二核苷酸的增加而形成一个完整的 TCTT 重复单位突变的结果，但突变机制还不清。另外 3 个较大等位基因（230，246，294bp）与小等位基因中观察到有规律 4bp 增加模式显著不同，表现出复杂的重复单位模式（表 13－34）。等位基因

在第 16 个重复单位含有一个非典型的 TCCT 四核苷酸；230bp 等位基因含有 17 个 TCTT 重复单位，在第 15 和 16 重复单位之间有 3 个非典型 CCTT 的插入，246bp 等位基因含有 21 个 TCTT 组成重复单位，在第 19 和 20 个重复单位之间被 3 个非典型 CCTT 四核苷酸隔开，294bp 等位基因的重复区更加复杂，这个等位基因有 28 个典型 TCTT 重复单位，在第 12 和 13 之间有 5 个 TCTG 四核苷酸，在第 26 和 27 之间有 3 个 CCTT；另外 230、246 和 294bp 的等位基因在第一个 TCTT 重复单位前没有 TC 二核苷酸的存在。表 13－35 为等位基因频率。

表 13－34　FGA 基因座等位基因重复区的序列与旁侧序列

等位基因	大小（bp）	重复序列
15	168	$(TTTC)_3$ TTTT TTCT $(CTTT)_7$ CTCC $(TTCC)_2$
17	176	$(TTTC)_3$ TTTT TTCT $(CTTT)_9$ CTCC $(TTCC)_2$
18	180	$(TTTC)_3$ TTTT TTCT $(CTTT)_{10}$ CTCC $(TTCC)_2$
18.2	182	$(TTTC)_3$ TTTT TT $(CTTT)_{11}$ CTCC $(TTCC)_2$
19	184	$(TTTC)_3$ TTTT TTCT $(CTTT)_{11}$ CTCC $(TTCC)_2$
20	188	$(TTTC)_3$ TTTT TTCT $(CTTT)_{12}$ CTCC $(TTCC)_2$
20.2	190	$(TTTC)_3$ TTTT TT $(CTTT)_{13}$ CTCC $(TTCC)_2$
21	192	$(TTTC)_3$ TTTT TTCT $(CTTT)_{13}$ CTCC $(TTCC)_2$
22	196	$(TTTC)_3$ TTTT TTCT $(CTTT)_{14}$ CTCC $(TTCC)_2$
22.2	198	$(TTTC)_3$ TTTT TT $(CTTT)_{15}$ CTCC $(TTCC)_2$
23	200	$(TTTC)_3$ TTTT TTCT $(CTTT)_{15}$ CTCC $(TTCC)_2$
24	204	$(TTTC)_3$ TTTT TT $(CTTT)_{16}$ CTCC $(TTCC)_2$
24.2	206	$(TTTC)_3$ TTTT TTCT $(CTTT)_{16}$ CTCC $(TTCC)_2$
25	208	$(TTTC)_3$ TTTT TTCT $(CTTT)_{17}$ CTCC $(TTCC)_2$
26	212	$(TTTC)_3$ TTTT TTCT $(CTTT)_{18}$ CTCC $(TTCC)_2$
27	216	$(TTTC)_3$ TTTT TTCT $(CTTT)_{19}$ CTCC $(TTCC)_2$
28	220	$(TTTC)_3$ TTTT TTCT $(CTTT)_{20}$ CTCC $(TTCC)_2$
29	224	$(TTTC)_3$ TTTT TTCT $(CTTT)_{15}$ CCTT $(CTTT)_5$ CTCC $(TTCC)_2$
34.2	264	$(TTTC)_4$ TTTT TT $(CTTT)_{18}$ $(CTTC)_3$ $(CTTT)_3$ CTCC $(TTCC)_4$
46.2	294	$(TTTC)_4$ TTTT TT $(CTTT)_{12}$ $(CTGT)_5$ $(CTTT)_{13}$ $(CTTC)_3$ $(CTTT)_3$ CTCC $(TTCC)_4$

旁侧序列：

等位基因 168－224bp（重复区）TCTC $(CTTC)_2$CTTT CTTC CTTT CTTTTTT（引物）3′

等位基因 230，246，294bp（重复区）TCTC $(CTTC)_4$CTTT CTTC CTTT CTTTTTT（引物）3′

表 13－35　Hum FGA 基因座等位基因频率

等位基因	中国汉族	中国汉族	中国汉族（北京）	中国壮族	美国白人	美国黑人
16	0.000			0		
16.2	0.000				0.000	0.0026
17	0.000			0	0.000	0.0026

续表

18	0.0578	0.02	0.0117	0	0.0150	0.0051
19	0.1111	0.035	0.0367	0.0049	0.0625	0.0462
19.2	0.000		0.0017		0.000	0.0026
20	0.0578	0.05	0.0333	0.0146	0.1625	0.0436
20.2	0.000	0.0050	0.0075	0.000		
21	0.2178	0.1017	0.1267	0.0437	0.1775	0.1333
21.2	0.000	0.0033	0.0100	0.0097	0.000	0.0026
22	0.3289	0.15	0.1750	0.0680	0.1650	0.1897
22.2	0.0178	0.001	0.0117	0	0.0050	0.0026
23	0.2089	0.2483	0.227	0.1553	0.1400	0.1897
23.2	0.0089	0.0133	0.015	0.0097	0.000	0.000
24	0.1956	0.1817	0.1933	0.1445	0.1325	0.1641
24.2	0.0044	0.0067	0.0133	0.0049	0.000	0.1641
25	0.1156	0.1083	0.0900	0.1408	0.1125	0.1231
25.2	0.0022	0.0083	0.0033	0.0049	0.000	0.000
26	0.0511	0.0450	0.0400	0.2039	0.0150	0.0410
26.2		0.0050				
27	0.0111	0.0083	0.0100	0.0922	0.0410	
28	0.0022	0.0017	0.0017	0.0485	0.0077	
29				0.0146		0.0026
杂合度	0.8400	0.8433	0.8552	0.868		
DP	0.9562	0.9416	0.9828	0.912		
PIC		0.8374	0.8346			
PE		0.7179		0.9026		

1. 引物序列　5′ – CCA TAG GTT TTG AAC TCA CAG – 3′

　　5′ – CTT CT CAG ATC CTC TGA CAC – 3′

2. PCR 反应体系　30μl 反应体系，含 50mmol/L KCl，10mmol/L Tris – HCl（pH 8.3），1.5mmol/L $MgCl_2$，dNTPs 200μmol/L，0.5u Taq DNA 聚合酶，5～10ng 模板 DNA，0.15μmol/L 引物。

3. 热循环参数　95℃2min 后，94℃60s，60℃60s，72℃60s，循环 30 次。72℃10min。

13.45 HumFOLP23 基因座

HUM FOLP23 基因座定位于人 6 号染色体，核心序列为 AAAC，已检出 7 个等位基因，片段长度范围为 162～184bp。表 13 – 36 为等位基因频率。

表 13-36　HUM FOLP23 基因座等位基因频率

等位基因*	中国汉族频率
n1	0.1277
n2	0.4946
n3	0.1902
n4	0.0516
n5	0.0163
n6	0.1168
n7	0.0027
杂合度	0.6882
DP	0.9259
PIC	0.6495
PE	0.4649

*根据片段大小命名等位基因

1. 引物序列　5′-ATT GTA AGA CTT TTG GAG CCATTT-3′

5′-TTC AGG GAG AAT GAG ATG GGC-3′

2. PCR 反应体系　20μl 反应体系，含有 50mmol/L KCl，10mmol/L Tris-HCl（pH 8.3），1.5mmol/L $MgCl_2$，0.001%白明胶，2.5mmol/L dNTPs 2μl，0.04μmol/L 引物，8μl Chelex-100 提取 DNA 溶液，0.5u Taq DNA 聚合酶 1。

3. 热循环参数　94℃ 1min，60℃1min，70℃1.5min，10 个循环；90℃1min，60℃1min，70℃1.5min，20 个循环，60℃30min。

13.46 HUMF13A01 基因座

HUMF13A01（F13A01）基因座是人凝血因子 X Ⅲ 的 A 亚基基因（human coagulation factor X Ⅲ a subunit gene），定位于人染色体 6p24-p25，（GenBank No.M21986），重复序列为 AAAC，已检出 16 个等位基因（表 13-37），片段长度为 177~235bp。

表 13-37　F13A01 基因座等位基因频率

等位基因	频率							
	中国汉族（北京）	中国汉族（太原）	中国汉族（广州）	中国汉族（台湾）	中国壮族（广西）	日本人	美国白人	美国黑人
2.2						0	0	0
3.2	0.3100	0.3238	0.2835	0.295	0.2765	0.328	0.085	0.087
4	0.1725	0.1666	0.1339	0.125	0.1810	0.098	0.041	0.076
5	0.0800	0.0714	0.0669	0.135	0.0524	0.028	0.208	0.342
6	0.4100	0.4285	0.4882	0.445	0.4857	0.543	0.287	0.131
7	0.0050	0.0095	0.0039			0.001	0.329	0.195
8	0.0225		0.0236				0.017	0.067
9								0.009
10								0.005

续表

11							0.009
12					0.001	0.002	0.011
13						0.005	0.032
14				0.0048		0.017	0.021
15						0.010	0.014
16							0.002
17							
杂合度	0.6851	0.79	0.61	0.6544	0.62	0.763	0.73
DP	0.8120	0.8305	0.844	0.8205			
PM		0.1695		0.1356			

1. 引物序列 5′－GAG GTT GCA CTC GAG CCT TTG CAA－3′

5′－TTC CTG AAT CAT CCC AGA GCC ACA －3′

或 5′－GAG GTT GCA CTC CAG CCT TT－3′

5′－ATGCCA TGC ACA TTA GAA A－3′

2. PCR 反应体系 在 25μl 反应体系，含有 50mmol/L KCl，10mmol/L Tris－HCl（pH 9.0），1.5mmol/L $MgCl_2$，0.001%白明胶，200μmol/L dNTPs，1.5pmol/L 引物，1u Taq DNA 聚合酶，5～15ng 模板 DNA。

3. 热循环参数 95℃ 1min，60℃45s，72℃ 2min，30 个循环。

13.47 HUMF13B 基因座（FⅩⅢ B）

HUMF13B 基因座是人凝血因子ⅩⅢ 的 b 亚基基因（human factor ⅩⅢ b subunit gene），又称 FⅩⅢ 基因座，核心序列 AAAT，定位于人类染色体 1q31－q32.1，已检出 6 个等位基因，片段长度为 169～185bp。表 13－38 为 F13B 基因座的等位基因频率。

表 13－38 F13B 基因座的等位基因频率

等位基因	频率				
	中国汉族（台湾）	澳大利亚人	美国白人	美国黑人	日本人
6	0.005	0.090	0.100	0.384	
7	0.000	0.016	0.020	0.157	0.003
8	0.045	0.245	0.259	0.100	0.065
9	0.225	0.225	0.215	0.243	0.203
10	0.725	0.419	0.0402	0.114	0.725
11		0.005	0.002	0.002	0.004
12			0.002		
杂合度	0.44	0.6991	0.6512	0.7318	0.43
DP	0.61	0.861			0.63
PIC		0.656			

1. 引物序列 5′－TGA GGT GGT GTA CTA CCA T－3′

5′ - GAT ATG AAA TTG CAC TCT A - 3′

2. PCR 反应体系　在 25μl 反应体系中，含有 50mmol/L KCl，10mmol/L Tris - HCl (pH8.8)，1.5mmol/L $MgCl_2$，0.1% Triton X - 100，200μmol/L dNTPs，0.5u Taq DNA 聚合酶，2ng 模板 DNA。

3. 热循环参数　94℃ 1min，60℃ 1min，72℃ 1.5min，30 个循环。

13.48 HUMLPL 基因座

HUMLPL 基因座位于人脂蛋白脂酶基因（lipoprotein lipase）第 6 内含子，又叫 HUMLIPOL 定位于人染色体 8q22，重复序列为 TTTA，已检出 6 个等位基因（表 13 - 39）片段长度为 125 ~ 175bp。

表 13 - 39　HUMLPL 基因座等位基因频率

等位基因	频		率	
	中国汉族	日本人	美国黑人	美国白人
7		0.003		
8				
9	0.01	0.001	0.002	0.002
10	0.675	0.712	0.146	0.047
11	0.1	0.101	0.370	0.412
12	0.19	0.180	0.151	0.287
13	0.025	0.003	0.272	0.203
14			0.059	0.049
杂合度	0.7342	0.45	0.69	0.77
DP	0.7	0.65		

1. 引物序列　5′ - CTG ACC AAG GAT AGT GGG ATA TAG - 3′
5′ - GGT AAC TGA GCG AGA CTG TGT CT - 3′
或　5′ - AAA TCT GAC CAA GGA TAG TGG G - 3′
5′ - CGA GTG ACG TTG GAG ACG AAA G - 3′

2. PCR 反应体系　20μl 反应液，含有 2μl 10 × PCR 缓冲液，0.2mmol/LdNTP，4μmol/L 引物各 2μl，0.5u Taq DNA 聚合酶，2μl 模板 DNA。

3. 热循环参数　96℃ 2min 后，94℃ 1min，60℃ 1min，70℃ 1.5min，10 个循环；90℃ 1min，60℃ 1min，70℃ 1.5min，20 个循环，60℃ 10min。

13.49 HUMMBP 基因座

HUMMBP 基因座是髓磷脂碱性蛋白（myelin basic protein）基因（GDB：J04746），定位于人染色体 18q23 - pter，由 Boylan（1990）最早报道的，含有二个 STR 系统，核心序列为 TGGA，由限制性内切酶 NlaⅢ 切割分离两个具有多态性体系 MBP - A，MBP - B。用一对引物可以扩增出 2 个内含子多态区：基因座 A（121 ~ 145bp）和基因座 B（209 ~ 237bp）。在 HUMMBP - B 基因座重复单位上游第 8 个核苷酸位置存在 G→A，导致产生无效等位基

因（图 13－3）。MBP－A 和 MBP－B 基因频率分布见表 13－40 和表 13－41。

HUMMBP 基因座 5′－旁侧序列

下划线部分为引物序列，引物 1 和 3 为基因座 A，引物 1 和 2 为基因座 B

CGAGGACAA GGACCTCGT AATTACAATCACTGTTAACTCAGATATTTAGAATGGT
1

A
2
CATTACACCTATTTGGTACTCATAAATAA (TGGA)$_n$TGGATGAACAG TAGGTAAA T
G

3
AGATGGATACATCAG（TGGA）TGAATCGATAGATGCATCAACAGCTAGGTAAA TGGAT

GAATACATGAGT

图 13－3 MBP 基因座序列

下划线部分为引物序列，引物 1 和 3 为基因座 A，引物 1 和 2 为基因座 B

表 13－40 HUMMBP－A 基因座等位基因频率

等位基因	频率	
	英国人	德国人
2	0	0.003
3	0	0.066
4	0.07	0.049
5	0.32	0.28
6	0.19	0.01
7	0.08	0.035
8	0.20	0.472
9	0.04	0.052
10	0.10	0.028
11	0	0.003
杂合度		0.72
DP		0.85

表 13－41 HUMMBP－B 基因座

等位基因	频率			
	德国人	美国人	西班牙人	丹麦人
7	0.374	0.26	0.330	0.3575
7C	0.003			
8		0.01		
9	0.038	0.13	0.045	0.0600
10	0.073	0.14	0.156	0.0725
10G→A				0.0225
11	0.339	0.25	0.326	0.3275
12	0.164	0.21	0.134	0.1600
13	0.007	0.009		
杂合度	0.70		0.71	
DP	0.80			

1. 引物序列　5′ - GAA CCT CGT GAA TTA CAA TC - 3′

5′ - ATT TAC CTA CCT GTT CAT CC - 3′

或用另一对引物　5′ - GAA CCT CGT GAA TTA CAA TC - 3′

5′ - CTC ATG TAT CCA TCT ATT TAC C - 3′

2. PCR 反应体系　25μl 反应体系，含有 1ng 模板 DNA，1u Taq DNA 聚合酶，1μmol/L 引物，200μmol/L dNTPs，2μl 10 × PCR 缓冲液

3. PCR 热循环参数　94℃30s，64℃30s，72℃1min，31 个循环。若采用第二对引物，则为 94℃30s，58℃30s，72℃1min，30 个循环。

13.50 HUMPLA2A 基因座

HUMPLA2A 基因座是胰磷脂酶 A - 2（pancreatic phospholipase A - 2）基因，定位于人染色体 12q23 - qter（GenBank Accession No. M22970），核心序列为 AAT，已检出 7 个等位基因，片段长度范围为 118 ~ 139bp。群体分布频率见表 13 - 42。

表 13 - 42　PLA2A 基因座等位基因频率

等位基因	频率	
	中国汉族（成都）	中国汉族（台湾）
10		0.005
11	0.2778	0.255
12	0.0093	0.015
13	0.1574	0.190
14	0.3055	0.305
15	0.1574	0.140
16	0.0833	0.085
17	0.0093	0.005
杂合度	0.79	0.79
DP	0.91	0.917

1. 引物序列　5′GGT TGT CTC CAT GAG GTT AGA - 3′

5′TTG AGC ACT TAC TAT GTG CCA GGC T - 3′

2. PCR 反应体系　20μl 反应体系，含有 2μl 10 × PCR 缓冲液，0.2mmol/L dNTPs，0.4μmol/L 引物，0.5u Taq DNA 聚合酶，模板 DNA 2μl。

3. 热循环参数　96℃ 2min 后，94℃1min，60℃1min，70℃1.5min，10 个循环，90℃ 1min，60℃1min，70℃1.5min，20 个循环，60℃10min。

13.51 HUMTH01 基因座

HUMTH01 基因座又叫 TC11 基因座（GenBank Accession No. D00269）是酪氨酸羟化酶（tyrosine hydroxylase）基因，定位于人染色体 11p15.5，重复单位为 AATG，已检出 8 个等位基因（等位基因 5 ~ 9，9.3，10，11），片段长度为 179 ~ 203bp。等位基因群体分布频率见表 13 - 43。

表 13－43 HUMTH01 基因座等位基因频率

等位基因	频率									
	中国汉族（北京）	中国汉族（上海）	中国汉族（青岛）	中国汉族（广州）	中国汉族（成都）	中国汉族（台湾）	中国壮族（广西）	中国苗族（云南）	中国藏族（拉萨）	日本人
5			0.0045	0.0039			0.0050			
6	0.0805	0.1436	0.1126	0.1047	0.232	0.142	0.0693	0.1087	0.097	0.242
7	0.286	0.2277	0.2387	0.2597	0.256	0.241	0.2475	0.2138	0.227	0.286
8	0.0508	0.0693	0.0631	0.0543	0.050	0.056	0.0693	0.0652	0.091	0.049
9	0.4894	0.3465	0.491	0.5233	0.475	0.483	0.3960	0.4928	0.481	0.382
9.3	0.0339	0.0891	0.0045	0.0155	0.033	0.009	0.0446	0.087	0.104	0.031
10	0.0539	0.1188	0.0856	0.0388	0.054	0.069	0.1634	0.0326		0.010
11		0.0049					0.0050			
杂合度	0.665	0.7268	0.6757	0.7502			0.7525	0.3677		0.70
DP	0.8319	0.9092	0.8572	0.8249		0.981	0.8899	0.8694		0.86

1. 引物序列 5′－GTG GGC TGA AAA GCT CCC GAT TAT－3′

5′－ATT CAA AGG GTA TCT GGG CTC TGG－3′

2. PCR 反应体系 25μl 反应体系，含有 1ng 模板 DNA，10mmol/L Tris－HCl（pH8.3），50mmol/L KCl，1.5mmol/L $MgCl_2$，0.1% Triton－X 100，0.001%白明胶，100μmol/L dNTPs，1u Taq DNA 聚合酶，0.3μmol/L 引物，。

3. 热循环参数 94℃20s，64℃20s，72℃1min，10 个循环，90℃20s，64℃20s，72℃1min，20 个循环。

13.52 HUMTPOX 基因座

HUMTPOX 基因座，又称 hTPO 基因座。是甲状腺过氧化物酶（thyroid peroxidase）基因，定位于人染色体 2p23－2pter，重复单位为 AATG，共有 7 个等位基因，片段长度为 224～252bp。群体等位基因分布频率见表 13－44。

表 13－44 TPOX 基因座等位基因频率

等位基因	频率						
	中国汉族（北京）	中国汉族（上海）	中国汉族（青岛）	中国汉族（台湾）	中国壮族（广西）	中国苗族（云南）	日本人
7	0.0173	0.0209		0.009	0.0050	0	
8	0.4926	0.4607	0.4775	0.543	0.5050	0.4348	0.451
9	0.151	0.1309	0.1351	0.099	0.1485	0.1014	0.121
10	0.052	0.0419	0.0315	0.013	0.0545	0.0217	0.036
11	0.2698	0.3141	0.3018	0.332	0.2723	0.3877	0.349
12	0.005	0.0314	0.0405	0.009	0.0149	0.0471	0.039
14	0.0124		0.0045	0.004		0.0072	0.004
杂合度	0.6618	0.6324	0.6847	0.621	0.5644	0.6087	0.65
DP	0.8278	0.8053	0.8297		0.8121	0.8112	0.82

1. 引物序列　5′ - ACT GGC ACA GAA CAG GCA CTT AGG - 3′
　　　　　　5′ - GGA GGA ACT GGG AAC CAC ACA GGT - 3′

2. PCR反应体系　25μl反应体系，含有10ng模板DNA，0.25u Taq DNA聚合酶，0.7μmol/L引物，200μmol/L dNTPs，7.5mmol/L Tris - HCl（pH 9.0），2.0mmol/L $(NH_4)_2SO_4$，0.01% Tween - 20，1.5mmol/L $MgCl_2$。

3. 热循环参数　96℃ 2min，94℃ 1min，64℃ 1min，72℃ 1.5min，30个循环。72℃ 10min。

13.53 HUMVWA基因座

冯．威布朗因子（Von Willebrand Factor，VWF）基因内含子40含有3个不同TCTA串联重复区，一个在1 900～2 000nt，一个在2 200～2 350nt，另一个在1 640～1 794nt。

HUMVWA基因座是冯．威布朗因子（Von Willebrand）基因因子内含子位于1640～1794nt的一个串联重复区，定位于人染色体12p - 12pter（GenBank accession No.M25858），也叫HUMVWA31/A基因，核心序列为TCTA和TCTG，已检出10个等位基因，片段长度为134～170bp，所有等位基因基本结构为TCTA $(TCTG)_{3-4}$（TCTA）n，但等位基因14例外，结构不同（表13－45）。表13－46为等位基因频率。

表13－45　HUMVWA基因座序列

等位基因命名	片段长度	序列结构
13	134	TCTA $(TCTG)_4$ $(TCTA)_8$
14	138	TCTA TCTG $(TCTG)_4$ $(TCTA)_3$TCCA $(TCTA)_3$
15	142	TCTA $(TCTG)_4$ $(TCTA)_{10}$
15′	142	TCTA $(TCTG)_3$ $(TCTA)_{11}$
16	146	TCTA $(TCTG)_4$ $(TCTA)_{11}$
16′	146	TCTA $(TCTG)_3$ $(TCTA)_{12}$
17	150	TCTA $(TCTG)_4$ $(TCTA)_{12}$
18	154	TCTA $(TCTG)_4$ $(TCTA)_{13}$
19	158	TCTA $(TCTG)_4$ $(TCTA)_{14}$
20	162	TCTA $(TCTG)_4$ $(TCTA)_{15}$
21	166	TCTA $(TCTG)_4$ $(TCTA)_{16}$
22	170	TCTA $(TCTG)_4$ $(TCTA)_{17}$

表13－46　HUMVWA基因座等位基因频率

等位基因	频率								
	中国汉族	中国汉族	中国汉族（北京）	中国汉族（成都）	中国汉族（太原）	中国汉族（广州）	中国壮族（广西）	中国藏族（西藏）	德国人
13			0.0150	0.004		0.0038			0.005
14	0.2273	0.2817	0.2625	0.232	0.2619	0.3030	0.3048	0.101	0.115
15	0.0411	0.0300	0.0375	0.012	0.0143	0.0227	0.0286	0.034	0.090
16	0.1515	0.0200	0.1650	0.157	0.2048	0.1932	0.1238	0.208	0.140
17	0.2576	0.2050	0.2425	0.302	0.2000	0.1818	0.2048	0.303	0.310

续表

18	0.1905	0.1997	0.1700	0.182	0.2048	0.1932	0.2143	0.208	0.225
19	0.1104	0.0750	0.0925	0.099	0.1095	0.0985	0.1048	0.129	0.095
20	0.0216	0.0116	0.0150	0.012	0.0048	0.0038	0.0095	0.017	0.015
>20							0.0095		0.005
H	0.8701	0.7934	0.8078		0.8029	0.9031	0.7957		
DP	0.9367	0.9301	0.9344		0.9290	0.9244	0.9239		
CE									
Pm					0.0710	0.0756			

1.引物序列 5′-CCC TAG TGG ATA AGA ATA ATC-3′

5′-GGA CAG ATG ATA AAT ACA TAG GAT GGA TGG-3′

2.PCR 反应体系 在 25μl 反应体系中，含有 1~20ng 模板 DNA，10mmol/L Tris-HCl (pH8.3)，50mmol/L KCl，1.5mmol/L $MgCl_2$，0.1% Triton-X 100，0.001%白明胶，1u Taq DNA 聚合酶，0.3μmol/L 引物，200μmol/L dNTPs。

3.热循环参数 94℃ 1min，54℃或 58℃ 1min，72℃ 1.5min，30 个循环。

13.54 vWF Ⅲ基因座

vWF Ⅲ基因座是位于 von willebrand 基因第 40 内含子中位于 1900~2000nt 的一个 STR 基因座，定位与人类第 12 号染色体短臂。已检出 9 个等位基因（表 13-47），基因长度为 154~182bp，中国汉族人群等位基因杂合度为 0.78，DP 0.921，PE 0.581。等位基因 11 和 12 频率为最高。这可能是由于 vWFⅢ多态性起源于等位基因 11 或 12，在人类进化过程中，由于滑动复制导致重复单位数目增加或减少而逐渐形成现在这样的群体多态性。

表 13-47 vWFⅢ基因座的等位基因群体频率

等位基因	频率		
	中国汉族	日本人	白种人
7		0.002	
8			
9	0.090	0.038	0.19
10	0.040	0.029	0.05
11	0.333	0.312	0.35
12	0.248	0.295	0.27
13	0.093	0.157	0.06
14	0.170	0.136	0.08
15	0.025	0.031	
16	0.003		
杂合度	0.78		
DP	0.921		

1. 引物序列　5′－TGT ACC TAG TTA TCT ATC CTG－3′

5′－GTG ATG ATG ATG GAG ACA G－3′

2. PCR 反应体系　25μl 反应体系，含有 10ng 模板 DNA，2.0mmol/L $MgCl_2$；0.2μmol/L 引物，200μmol/L dNTPs，50mmol/L KCl，10mmol/L Tris－HCl（pH 9.0），1% Triton X－100，1u Taq DNA 聚合酶。

3. 热循环参数　94℃45s，58℃1.5min，72℃1.5min，30 个循环。

第十四章 性染色体 STR 基因座

14.1 Y 染色体 STR 基因座

Y 染色体属人类的性染色体，只有正常男性拥有，女性没有。以往 Y 染色体一直被用于性别鉴定，直到 20 世纪 90 年代中后期，对 Y 染色体特有的 STR 序列及多态性进行了较深入的研究，开始选用一些 Y 染色体上的 STR 基因座进行个人识别和亲子鉴定。有些 Y-STR 基因座已被列入常规法医检验项目。

14.1.1 Y 染色体 STR 基因座特性

Y 染色体除拟常染色质区外，大部分在减数分裂时不会出现重组交换。Y 染色体特异基因座上的等位基因是单倍体，序列结构能够稳定地由父亲传递给儿子。Y 染色体的这种特性在个人识别、父权鉴定中有一定意义，利用 Y 染色体遗传特征还能追溯父系的迁移历史及重新构建同一父系家族。所有 Y-STR 基因座位于一条染色体上，不论它们的物理距离多远，均是连锁遗传，个体识别与父权鉴定相关机率计算不能参照常染色体的累计法，只能以单倍型概率计算。

Y-STR 基因座的变异、突变率一般不超过常染色体突变率，即使高度多态性的 DYS385 和 CYAⅡ基因座的突变率也在常染色体突变率范围内。与常染色体 STR 相比，Y-STR 更经常发生"无效等位基因"或多个等位基因。如在一个父子对中观察到三个 DYS19 等位基因，在 2 个个体中观察到了 DYS19 的二重等位基因，在一个个体中观察到 DYS390 的二重等位基因，在一个个体中观察到 DYS385 的三重等位基因等。这些都可能是由于 Y 染色体的特定结构所引起。细胞遗传分析发现在 Y 染色体长臂上的异源染色区出现长度变化现象。

14.1.2 Y 染色体 STR 等位基因的命名

Y 染色体特异 STR 基因座等位基因名称按国际法医血液遗传学会 DNA 委员会（DNA Commission of the ISFH，1994）推荐的原则命名，并在 1996 年在柏林召开了第一届 Y 染色体 STR 分型的学术会议，统一了 Y-STR 基因座等位基因的命名原则，用可变重复序列的数目命名 Y-STR 基因座等位基因。对暂时还没有完全测出等位基因序列的或还没有确定等位基因长度的那些 Y-STRs 基因座，可直接用等位基因的长度命名，或用简单的数字序号表示，以观察到的最短的等位基因命名为等位基因 1。

14.2 常用的 Y 染色体 STR 基因座

14.2.1 DYS19 基因座

DYS19 基因座（GDB：121409），又称 Y27H39，位于 Y 染色体的短臂，重复序列为四核苷酸，10～19 个（CTAT/C）重复，已观察到 10 个等位基因。常见等位基因片段大小为

186～206bp，群体基因差异度在 0.56～0.698。等位基因序列结构为 P1（21bp）－8bp $(at)_5$agtatt $(at)_4$agtgtt $(at)_5$agtgttt $(TAGA)_3$tagg $(TAGA)_n$ tata31bp－p2（19bp），群体分布频率见表 14－1。

表 14－1　DYS19 基因座群体基因频率

等位基因	中国汉族（东北）	中国汉族（广州）	中国汉族（沈阳）	中国汉族（成都）	中国汉族（成都）	日本人
13（186）	0.029	0.189	0.045	0.0366	0.069	0.054
14（190）	0.2609	0.360	0.2342	0.1707	0.2644	0.071
15（194）	0.2609	0.360	0.3694	0.5976	0.3908	0.457
16（198）	0.2998	0.072	0.2703	0.1463	0.2069	0.190
17（202）	0.1594	0.018	0.0811	0.0366	0.0575	0.223
18（206）				0.0122	0.0115	0.005
19（210）						
GD 值		0.698				0.697

1. 引物序列　5′－CTA CTG AGT TTC TGT TAT AGT－3′

5′－ATG GCA TGT AGT GAG GAC A－3′

2. 反应体系　25μl 反应体系中，含有 1u Taq DNA 聚合酶，0.25μmol/L 每种引物，200μmol/L dNTPs，1.5mmol/L $MgCl_2$，50mmol/L KCl，10mmol/L Tris HCl（pH 8.3）。

3. 扩增循环参数　94℃ 3min 1 个循环，94℃30s，51℃30s，72℃ 90s，30 个循环；或 94℃ 1min，56 ℃1min，72℃ 1min，29 个循环。

14.2.2 DYS385 **基因座**

DYS385 基因座（GDB：316257）重复序列为四核苷酸，引物有二个结合部位，可以扩增出两个大小不同的等位基因片段，片段大小在 357～405bp 之间。但是此两个区的扩增片段大小互相交叉，相互间无法区分。因此必须用单倍体型来表示，已观察到 13 种片段。群体分布频率见表 14－2。

表 14－2　DYS385 单倍型组合群体分布频率

单倍型	德国人萨克森	意大利人	西班牙人	单倍型	德国人萨克森	意大利人	西班牙人
9/13			0.0108	12/12		0.016	0.043
9/14	0.004			12/13	0.004	0.016	0.0215
10/14	0.068	0.008	0.0108	12/14	0.052	0.048	0.0645
10/15	0.02			12/15	0.008	0.024	0.0645
10/16	0.008			12/16	0.004	0.008	
10/18		0.008		12/17	0.004	0.016	0.0215
11/11		0.04		13/13	0.008	0.016	0.0215
11/12			0.0108	13/14	0.048	0.016	0.0322
11/13	0.036		0.0322	13/15	0.036	0.04	0.0860
11/14	0.412	0.16	0.3226	13/16	0.004	0.04	0.0215
11/15	0.068	0.04	0.0215				
11/16	0.016	0.04	0.0430				

续表

11/17		0.016					
13/17	0.02	0.064	0.0108	15/19	0.004		
13/18		0.008		16/16	0.004	0.024	0.0108
13/19		0.04		16/17	0.008	0.016	
13/20		0.008		16/18	0.028	0.032	0.0108
14/14	0.024	0.016	0.0322	16/19	0.004	0.04	
14/15	0.044	0.04	0.0215	16/21		0.008	
14/16	0.008	0.048	0.0430	17/17		0.008	
14/17		0.024	0.0215	17/18	0.016	0.016	
14/18		0.016	0.0108	17/19	0.004	0.008	
15/15	0.008	0.016	0.0322				
15/16	0.020	0.024					
15/17	0.008						
15/18		0.008					
				单倍型GD值	0.808		0.8669

1. 引物序列　5′-AGC ATG GGT GAC AGA GCT A-3′

5′-GGG ATG CTA GGT AAA GCT G-3′

2. PCR 反应体系　25μl 反应体系内含 1u Taq DNA 聚合酶，0.1～0.8 μmol/L 每种引物，200μmol/L dNTPs，1.5mmol/L $MgCl_2$，50mmol/L KCl，10mmol/L Tris HCl pH 8.3。

3. 扩增循环参数　94℃3min 1 个循环，94℃60s，55℃60s，72℃90s，30 个循环。

14.2.3 DYS389 基因座

DYS389 基因座含有 TCTG 和 TCTA 二种重复单位的二个重复区，所有等位基因结构为 $(TCTG)_n$ $(TCTA)_m$ (48bp) $(TCTG)_p$ $(TCTA)_q$，n=2～7，m=9～15，p=3，q=8～13。在重复单位 $(TCTG)_n$和 $(TCTG)_p$ 的5′端均含有一个共同固定序列，使一个 PCR 引物有二个退火结合基因座，可以扩增出二个大小不同的片段，因此，可将 DYS389 基因座分为 DYS389Ⅰ和 DYS389Ⅱ基因座。DYS389Ⅰ基因座含有（TCTG）p（TCTA）q，其等位基因片段比 DYS389Ⅱ基因座小 100bp 以上。DYS389Ⅱ基因座色含 DYS389Ⅰ序列，其全序列为 $(TCTG)_n$ $(TCTA)_m$ (48bp) $(TCTG)_p$ $(TCTA)_q$。

1. DYS389Ⅰ基因座

DYS389Ⅰ基因座（GDB：366108）重复序列为四核苷酸 TATC 或表达为 TCTA，已观察到 5 个等位基因，等位基因片段在 243～259bp 之间。等位基因序列结构为 P1（25bp）TATC $(TGTC)_3$ $(TATC)_n$cctccctcTATCaatcTATCtattTATCtagc71bpTGTC43bp-p2（20bp），群体分布频率见表 14-3。

表 14－3　DYS389 Ⅰ 群体分布频率

等位基因	频		率
	中国汉族（广州）	中国汉族（沈阳）	中国维族（新疆）
8（243）	0.063	0.0159	
9（247）	0.522	0.4127	0.268
10（251）	0.171	0.4444	0.446
11（255）	0.243	0.1270	0.286
12（259）			

分型方法见 DYS389 Ⅱ 的分型方法。

2.DYS389 Ⅱ 基因座

DYS389Ⅱ基因座（GDB：366108）重复序列为四核苷酸，已观察到 7 个等位基因，等位基因片断大小为 359～383bp。Wolf.B（1998）用测序法及半巢式 PCR（nested－PCR）法测定重复单位具体的重复数目。其等位基因序列结构见表 14－4，群体分布频率见表 14－5。

表 14－4　DYS389 Ⅱ 基因座序列结构（Pestoni C et al，1998）

共同结构	
	P1（25bp）－6bp（TCTG)$_n$（TCTA)$_m$48bp（TCTG)$_3$（TCTA)$_p$152bp－p2（20bp）
等位基因（bp）	序列
24（359）：	P1（25bp）－6bp（TCTG)$_5$（TCTA)$_{10}$48bp（TCTG)$_3$（TCTA)$_9$152bp－p2（20bp）
25（363）：	P1（25bp）－6bp（TCTG)$_4$（TCTA)$_{12}$48bp －－－－－ （TCTA)$_9$152bp－p2（20bp）
26（367）：	P1（25bp）－6bp（TCTG)$_5$（TCTA)$_{11}$48bp －－－－－ （TCTA)$_{10}$ 152bp－p2（20bp）
27（371）：	P1（25bp）－6bp（TCTG)$_5$（TCTA)$_{11}$48bp －－－－－ （TCTA)$_{11}$ 152bp－p2（20bp）
28（375）：	P1（25bp）－6bp（TCTG)$_5$（TCTA)$_{11}$48bp －－－－－ （TCTA)$_{12}$ 152bp－p2（20bp）
29（379）：	P1（25bp）－6bp（TCTG)$_5$（TCTA)$_{14}$48bp －－－－－ （TCTA)$_{10}$ 152bp－p2（20bp）

表 14－5　DYS389 Ⅱ 单体型群体分布频率

等位基因	中国汉族（沈阳）	中国汉族（成都）	中国维族（新疆）
23			
24	0.0952		
25	0.1429		0.107
26	0.4286	0.069	0.339
27	0.2698	0.023	0.196
28	0.0635	0.3793	0.286
29		0.2414	0.071
30		0.2069	
31		0.0575	
32		0.023	

(1) 引物序列 5′-CCA ACT CTC ACT TGT ATT ATC TAT-3′

5′-TCT TAT CTC CAC CCA CCA GA-3′

(2) 反应体系 25μl 内含 1u Taq DNA 聚合酶，0.1~0.8μmol/L 每种引物，200μmol/L dNTPs，1.5mmol/L $MgCl_2$，50mmol/L KCl，10mmol/L Tris HCl（pH 8.3）。正链引物用荧光染料标记。

(3) PCR 扩增循环条件 94℃ 3min 1 个循环，94℃ 15s，58℃ 20s，72℃ 20s，5 个循环后，94℃ 15s，54℃20s，72℃20s，30 个循环。

14.2.4 DYS390 基因座

DYS390 基因座（GDB：366115）为四核苷酸重复序列，重复序列为 CTC/AT 或表达为 TCTG/A，已观察到 10 个等位基因，重复数目为 18~27，常见等位基因片段大小为 203~227bp，其常见等位基因序列结构见表 14-6，群体分布频率见表 14-7。

表 14-6 DYS390 基因座的测序数据（Pestoni C et al，1998）

共同结构
P1(23bp)-26bp(TCTA)2(TCTG)n(TCTA)mTCTG(TCTA)4TCA(TCTA)2 29bp-P2(20bp)
等位基因(bp) 序列
21(204):P1(23bp)-26bp$(TCTA)_2(TCTG)_8(TCTA)_8$TCTG$(TCTA)_4$TCA$(TCTA)_2$ 29bp-P2(20bp)
22(208):P1(23bp)-26bp-----$(TCTG)_8(TCTA)_9$-------------29bp-P2(20bp)
23(212):P1(23bp)-26bp-----$(TCTG)_8(TCTA)_{10}$-------------29bp-P2(20bp)
P1(23bp)-26bp-----$(TCTG)_9(TCTA)_9$-------------29bp-P2(20bp)
24(216):P1(23bp)-26bp-----$(TCTG)_8(TCTA)_{11}$-------------29bp-P2(20bp)
25(220):P1(23bp)-26bp-----$(TCTG)_8(TCTA)_{12}$-------------29bp-P2(20bp)

表 14-7 DYS390 基因座群体分布频率

等位基因	中国汉族(华北)	中国汉族(广州)	中国汉族(成都)	中国维族(新疆)	德国人
21	0.0103	0.027		0.018	0.008
22	0.0206	0.099	0.0366	0.054	0.084
23	0.3916	0.468	0.4878	0.196	0.256
24	0.3814	0.279	0.3292	0.429	0.328
25	0.1549	0.126	0.1098	0.286	0.304
26	0.0412	0.0366	0.018	0.02	

1. 引物序列 5′-TAT ATT TTA CAC ATT TTT GGG CC-3′

5′-TGA CAG TAA AAT GAA CAC ATT GC-3′

2. 反应体系 在 25μl 反应体系内，含 1u Taq DNA 聚合酶，0.1~0.8 μmol/L 每种引物，200 μmol/L dNTPs，1.5mmol/L $MgCl_2$，50mmol/L KCl，10mmol/L Tris HCl(pH 8.3)。

3. PCR 扩增循环参数 94℃ 变性 3min 后，94℃ 15s，58℃ 20s，72℃ 20s，5 个循环；94℃

15s,54℃20s,72℃20s,30 个循环。或先 94℃变性 3min,然后 94℃1min,55℃ 40s,70℃ 40s,循环 32 次。

14.2.5 DYS391 基因座

DYS391 基因座重复序列为四核苷酸,X、Y 染色体上均有此基因座的多态性,等位基因片段大小在 279～291bp 之间,已观察到 5 个等位基因。男性群体分布频率见表 14－8。

表 14－8　DYS391 群体分布频率

等位基因	中国汉族(沈阳)	中国汉族(成都)	中国汉族(成都)	泰国人	日本仙台人
6	0.009		0.024		
9	0.018	0.0805	0.0244	0.06	0.0446
10	0.7117	0.7126	0.6951	0.78	0.8929
11	0.2613	0.2069	0.2561	0.17	0.0536
12				0.02	0.0089

1. 引物序列　5′－GAT TTC AAT CAA TCA TAC ACC CA－3′

5′－GAT TCT TTG TGG TGG GTC TG－3′

2. 反应体系　在 25μl 反应体系内，含 1u Taq DNA 聚合酶，0.1～0.8 μmol/L 每种引物，200 μmol/L dNTPs，1.5mmol/L $MgCl_2$，50mmol/L KCl，10mmol/L Tris HCl (pH 8.3)。正链引物用荧光染料标记。

3. 扩增循环参数　94℃ 3min 1 个循环；94℃ 15s，58℃ 20s，72℃－20s，5 个循环；94℃ 15s，54℃20s，72℃20s，30 循环周期。或 94℃ 1min，56℃ 30s，72℃1min，28 个循环。

14.2.6 DYS392 基因座

DYS392 基因座重复序列为三核苷酸，7～16 个 ATT 重复，已发现 10 个等位基因，片段大小在 239～260bp 之间。群体分布频率见表 14－9。

表 14－9　DYS392 群体分布频率

等位基因	中国汉族（沈阳）	中国汉族（成都）	日本人	泰国人
7			0.005	
8			0.005	
9			0.000	
10	0.009	0.0115	0.085	
11	0.1622	0.0805	0.530	0.24
12	0.1351	0.0345	0.145	0.12
13	0.2973	0.2759	0.205	0.38
14	0.3333	0.5517	0.020	0.26
15	0.0541	0.0460	0.000	
16	0.009		0.005	
17				
GD 值			0.65	.

1. 引物序列 5′-TCA TTA ATC TAG CTT TTA AAA ACA A-3′

5′-AGA CCC AGT TGA TGC AAT GT-3′

2. 反应体系 在 25μl 反应体系内，含 1u Taq DNA 聚合酶，0.1~0.8 μmol/L 每种引物，200 μmol/L dNTPs，1.5mmol/L $MgCl_2$，50mmol/L KCl，10mmol/L Tris-HCl (pH 8.3)。

3. 热循环参数 94℃ 3min 1 个循环；94℃ 15s，58℃ 20s，72℃ 20s，5 个循环；94℃ 15s，54℃20s，72℃20s，30 个循环。

14.2.7 DYS393 基因座

DYS393 基因座（GDB：456649）为四核苷酸重复序列，等位基因片段长度在 120~132bp 之间，已观察到 6 个等位基因。其等位基因序列结构为 P1 (24bp) - $(AGAT)_n$ $(ATGT)_2$ 17bp-P2 (22bp)，群体分布频率见表 14-10。X 染色体上也有此基因座，而且 X、Y 染色体的扩增产物片段范围一致。

表 14-10 DYS393 基因座群体分布频率

等位基因	中国汉族（沈阳）	中国汉族（成都）	中国汉族（成都）	泰国人
10				
11 (115)	0.009	0.0115		
12 (119)	0.4414	0.4828	0.4268	0.28
13 (123)	0.3063	0.3333	0.2927	0.42
14 (127)	0.1712	0.1149	0.2073	0.24
15 (131)	0.0541	0.0575	0.0732	0.06
16 (135)	0.018			

1. 引物序列 5′-GTG GTC TTC TAC TTG TGT CAA TAC -3′

5′-AAC TCA AGT CCA AAA AAT GAG G-3′

2.PCR 反应体系 25μl 反应体系内含 1u Taq DNA 聚合酶，0.25 μmol/L 每种引物，200 μmol/L dNTPs，1.5mmol/L $MgCl_2$，50mmol/L KCl，10mmol/L Tris-HCl (pH 8.3)。

3. 热循环参数 94℃3min 1 个循环，94℃15s，58℃20s，72℃20s，30 个循环。

14.2.8 DXYS156Y 基因座

DXYS156Y 基因座重复序列为五核苷酸，等位基因长度在 150~260bp 之间。其等位基因群体分布频率见表 14-11。

表 14-11 DXYS156Y 等位基因群体分布频率

等位基因	泰国人	库尔德人
7		
8		
9		0.0198
10		
11	0.14	0.1287
12	0.56	0.8515
13	0.14	
14	0.16	

1. 引物序列　5′ - GTA GTG GTC TTT TGC CTC C - 3′

5′ - CAG ATA CCA AGG TGA GAA TC - 3′

2. PCR 反应体系　25μl 反应体系内含 1u Taq DNA 聚合酶，0.25μmol/L 每种引物，200 μmol/L dNTPs，1.5mmol/L $MgCl_2$，50mmol/L KCl，10mmol/L Tris - HCl（pH 8.3）。

3. 扩增循环参数　94℃变性 3min 后，94℃15s，58℃20s，72℃20s，30 个循环。

14.2.9 A10 基因座

A10（G42674）基因座是 1999 年发现的一个 Y - STR 基因座。其核心序列为单一的 TATC 的四核苷酸序列，对 PCR 扩增产物进行序列分析，发现 A10 基因座的基本结构为：

5′ - P1（28bp） - 46bp - （TATC)$_n$ - 16bp - P2（24bp） - 3′。其等位基因群体分布频率见表 14 - 12。

表 14 - 12　A10 基因座等位基因群体分布频率

等位基因	片段长度（bp）	频率	
		中国汉族	日本人
10	154		0.009
11	158	0.01	0.036
12	162	0.48	0.450
13	166	0.31	0.378
14	170	0.12	0.099
15	174	0.07	0.018
16	178	0.01	0.009
GD		0.6505	0.64

1. 引物序列　5′ - CCT GCC ATC TCT ATT TAT CTT GCA TAT A - 3′

5′ - ATA AAT GGA GAT AGT GGG TGG ATT - 3′

2. PCR 反应　在 10μl 反应体系中，含有 8mmol/L Tris - HCl（pH8.3），40mmol/L KCl，1.5mmol/L $MgCl_2$，0.16mmol/L dNTPs，800ng/ml BSA，1u 金牌 Taq DNA 聚合酶，0.2μmol/L 的每一种引物，1μl Chelex - 100 提取液。

3. 热循环参数　95℃变性 10min 后，94℃30s，58℃30s，72℃30 s，循环 5 次；94℃ 30s，56℃30s，72℃30s，循环 30 次，最后 60℃30min，4℃保存。

14.2.10 C4 基因座

C4（G42674）基因座是 1999 年发现的一个 Y - STR 基因座。C4 基因座的基本结构为：5′- P1（26bp） - 70bp - （TATC)$_3$（TATG)$_2$（TATC)$_2$（TATG)$_2$（TATC)$_n$ - 49bp - P2（29bp） - 3′。

重复序列虽有二种序列，但有变化的为单一的 TATC 的四核苷酸序列。其等位基因群体分布频率见表 14 - 13。

表 14-13 C4 基因座等位基因群体分布频率

等位基因	片段长度（bp）	频率	
		中国汉族	日本人
9	242		0.009
10	246	0.01	0.027
11	250	0.11	0.288
12	254	0.31	0.468
13	258	0.30	0.135
14	262	0.17	0.045
15	266	0.06	0.009
16	270	0.03	0.009
17	274		0.009
18	278	0.01	
GD		0.7659	0.68

1. 引物序列　5′-AGT GTC TCA CTT CAA GCA CCA AGC AC-3′

5′-GCA GCA AAA TTC ACA GTT GGA AAA ATG T-3′

2. PCR 反应　在 10μl 反应体系中，含有 8mmol/L Tris-HCl（pH8.3），40mmol/L KCl，1.5mmol/L $MgCl_2$，0.16mmol/L dNTPs，800ng/ml BSA，1u 金牌 Taq DNA 聚合酶，0.2μmol/L 的每一种引物，1μl Chelex-100 提取液。

3. 热循环参数　95℃10min 预变性后，94℃30s，58℃30s，72℃30s，循环 5 次；94℃30s，56℃30s，72℃30s，循环 30 次，最后 60℃30min，4℃保存。

14.2.11 A7.1 基因座

A7.1（G42674）基因座是 1999 年发现的一个 Y-STR 基因座。其核心序列为 GATA，等位基因片段长度为 161～181bp。其等位基因群体分布频率见表 14-14。

表 14-14 A7.1 基因座等位基因群体分布频率

等位基因	汉族	壮族
7	0.04	0.01
8	0.16	0.00
9	0.34	0.22
10	0.34	0.56
11	0.12	0.20
12	0.00	0.01
GD	0.7272	0.5978

1. 引物序列　5′-CAA ATT TGC CAA ACT CTT TC-3′

5′-TCT ATC CTC TGC CTA TCA TTT ATT A-3′

2. PCR 反应　在 10μl 反应体系中，含有 8mmol/L Tris-HCl（pH8.3），40mmol/L KCl，1.5mmol/L $MgCl_2$，0.16mmol/L dNTPs，800ng/ml BSA，1u 金牌 Taq DNA 聚合酶，0.2μmol/L 的每一种引物，1μl Chelex-100 提取液。

3. 热循环参数 95℃10min 后，94℃30s，56℃30s，72℃30 s，循环 35 次，最后 60℃ 45min，4℃保存变性。

14.2.12 A7.2 **基因座**

A7.2（G42671）基因座是 1999 年发现的一个 Y - STR 基因座。其核心序列为 GATA，等位基因片段长度为 161～181bp。其等位基因群体分布频率见表 14 - 15。

表 14 - 15 A7.1 基因座等位基因群体分布频率

等位基因	汉族	壮族
8	0.02	0.00
9	0.02	0.07
10	0.20	0.45
11	0.51	0.33
12	0.23	0.13
13	0.02	0.02
GD	0.6458	0.664

1. 引物序列 5′ - AGG CAG AGG ATA GAT GAT ATG GAT - 3′

 5′ - TTC AGG TAA ATC TGT CCA GTA GTGA - 3′

2. PCR 反应 在 10μl 反应体系中，含有 8mmol/L Tris - HCl（pH8.3），40mmol/L KCl，1.5mmol/L $MgCl_2$，0.16mmol/L dNTPs，800ng/ml BSA，1u 金牌 Taq DNA 聚合酶，0.2μmol/L 的每一种引物，1μl Chelex - 100 提取液。

3. 热循环参数 95℃变性 10min 后，94℃30s，56℃30s，72℃30 s，循环 35 次，最后 60℃45min，4℃保存。

14.2.13 A4（G42670）**基因座**

核心序列为 AGAT，扩增片段长度为 239～251bp，重复单位数为 19～22。

1. 引物 5′- TCG ACT CGG TAC CAC TTC CTA GGT TTT C - 3′

 5′- TGG CTT GGA ATT CTT TTA CCC ATC ATC T - 3′

2. PCR 扩增条件 同 A10 基因座。

14.2.14 H4（G42676）**基因座**

核心序列为 GATA，扩增片段长度为 356～372bp，重复单位数为 26～30。

1. 引物 5′- TCG ACT CGG TAC CAC TTC CTA GGT TTT C - 3′

 5′- TGG CTT GGA ATT CTT TTA CCC ATC ATC T - 3′

2. PCR 扩增条件 同 A10 基因座。

14.2.15 DYS434 **基因座**

核心序列为 ATCT，片段长度为 110～122bp，重复单位数为 8～11。

1. 引物 5′- CAC TCC CTG AGT GCT GGA TT - 3′

 5′- GGABGATBGAABTGA ATG GAT GGA - 3′

2. 反应体系 25μl 反应体系中，含有 1×PCR 缓冲液，200μmol/L dNTPs，1.5mmol/L $MgCl_2$，1u Gold Taq DNA 聚合酶，0.4μmol/L 每种引物，5ng DNA.

3. 循环参数 94℃10min 后，94℃30s，56℃30s，72℃30s，退循环 30 次，最后 60℃

30min。

14.2.16 DYS435 基因座

核心序列为 TGGA，片段长度范围为 218～222bp，重复数为 15～16。

1. 引物 5′－AGC ATC TCC ACA CAG CAC－3′

5′－TTC TCT CTC CCC CTC CTC TC－3′

2.PCR 扩增条件 同 DYS434 基因座。

14.2.17 DYS436 基因座

核心序列为 GTT，片段长度为 130～133bp。

1. 引物 5′－CCA GGA GAG CAC ACA CAA AA－3′

5′－GCA ATC CAA CTT CAG CCA AT－3′

2.PCR 扩增条件 同 DYS434 基因座。

14.2.18 DYS437 基因座

核心序列为 TCTA/TCTG，结构为（TCTA）n（TCTG）m（TCTA）423bp（TCTA）2，n＝8～11，m＝1～2，片段长度范围为 184～192bp。

1. 引物 5′－GAC TAT GGG CGT GAG TGC AT－3′

5′－AGA CCC TGT CAT TCA CAG ATG A－3′

2.PCR 扩增条件 同 DYS434 基因座。

14.2.19 DYS438 基因座

核心序列为 TTTTC，长度范围为 216～236bp，重复单位数为 9～13

1. 引物 5′－TGG GGA ATA GTT GAA CGG TAA－3′

5′－GTG GCA GAC GCC TAT AAT CC－3′

2.PCR 扩增条件 同 DYS434 基因座。

14.2.20 DYS439 基因座

核心序列为 AGAT，片段长度范围为 240～248bp，重复单位数为 19～21。

1. 引物 5′－TCC TGA ATG GTA CTT CCT AGG TTT－3′

5′－GCC TGG CTT GGA ATT CTT TT－3′

2.PCR 扩增条件 同 DYS434 基因座。

A7.1，A7.2，A10，C4，H4，DYS438 和 DYS439 基因座可以复合在一起组成 7 重复合扩增。在 25μl 反应体系中，含有 1×PCR 缓冲液，200μmol/L dNTPs，2.0mmol/L $MgCl_2$，2u Gold Taq DNA 聚合酶，5ng DNA，每个基因座的浓度分别是 1.8μmol/L（A7.1），0.08μmol/L（A7.2）基因座 0.12μmol/L（A10），0.08μmol/L（C4），0.12μmol/L（H4），0.6μmol/L（DYS438）和 0.32μmol/L（DYS439）。

循环参数 94℃10min 后，94℃1min，64℃1min，68℃2min，退火温度按 0.5℃/循环降低，循环 12 次。然后，94℃1min，58℃1min，68℃2min，，循环 18 次，最后 60℃30min。

DYS436，DYS435，DYS436 和 DYS437 基因座可以复合扩增。扩增反应条件同上述 6 重复合扩增体系，引物浓度分别是 0.2μmol/L（DYS434），0.05μmol/L（DYS435），0.05μmol/L（DYS436），0.2μmol/L（DYS437）。

14.3 常用 Y-STR 基因座单倍型（组）群体数据

由于 Y 染色体非拟常染区不与 X 染色体重组，Y 特有的 STR 基因座为父系遗传且相互连锁，以等位基因单倍型表达遗传差异性。在 82 名成都汉族个体的 5 个 Y-STR 基因座 DYS19（6 个等位基因）、DYS389（4 个等位基因）、DYS390（5 个等位基因）、DYS391（4 个等位基因）、DYS393（4 个等位基因）只观察到 58 种单倍型，有 47 种单倍型在所调查的人群中只出现一次，有 11 种单倍型至少出现二次或二次以上，出现频率最高（6/82）的只有一个单倍型（侯一平等 1999）。在 111 名广州地区汉族个体的 8 个 Y-STR 基因座 DYS19、DYS389Ⅰ/Ⅱ、DYS390、DYS391、DYS392、DYS393、DYS385 观察到 110 种单倍型（唐双柏等），同样地在 100 名意大利人的 8 个 Y-STR 基因座 DYS19（4 个等位基因）、DYS389Ⅰ（4 个等位基因）/Ⅱ（6 个等位基因）、DYS390（5 个等位基因）、DYS391（4 个等位基因）、DYS392（5 个等位基因）、DYS393（4 个等位基因）、DXYS156（2 个等位基因）观察到 71 种单倍型，有 15 种单倍型至少在 2 名以上个体存在，一个单倍型最多为 13 个个体共有（Rossi E，1998）。单倍型的这种群体分布情况见表 14-16，表 14-17。

表 14-16　Y-STR 基因座群体单倍型组合

群体	观察个体数	Y-STR 基因座单倍型	单倍型数	一个单倍型最多个体数	参考文献
汉族（广州）	111	DYS19-DYS389Ⅰ/Ⅱ-DYS390-DYS391-DYS392-DYS393-DYS385	110	2	唐双柏等(待发表)
汉族（成都）	82	DYS19-DYS389-DYS390-DYS391-DYS393	58	6	侯一平等(1999)
泰国人	50	DYS19-DSYS156Y-DYS389(m,n,p,q)-DYS390-DYS391-DYS392-DYS393	44	3	Horst B (1999)
库尔德人	101	DYS19-DSYS156Y-DYS389(m,n,p,q)-DYS390-DYS391-DYS392-DYS393	82	10	BrinkmannC (1999)
意大利人中南部	125	DYS19-DYS389Ⅰ/Ⅱ-DYS390-DYS391-DYS392-DYS393-DYS385	117	2	CagliaA(1998)
意大利人 Modena 区	100	DYS19-DYS389Ⅰ/Ⅱ-DYS390-DYS391-DYS392-DYS393-DXYS156	71	13	Rossi E(1998)
德国人 Saxony	250	DYS19-DYS390-DYS389Ⅰ/Ⅱ-DYS385	164	16	Lessig R(1998)
德国人波恩地区	102	DYS19-DYS389Ⅰ/Ⅱ-DYS390-DYS3393	56	13	Junge A(1998)

表 14-17　7 个基因座在 9 个人群 Y 染色体单倍型资料

	中国汉族	蒙古人(Khalkh)	荷兰人	德国人	巴斯克人	奥地利人	Inuits(格陵兰)	苏里南人	俾格米人	Comb
个体数	36	40	88	70	30	63	62	54	31	474
单倍型数	34	29	65	63	17	40	22	47	18	301
识别率%	94.4	72.5	73.9	90.0	56.7	63.5	35.5	87.0	58.1	-
单倍体差异性	0.99	0.97	0.98	0.99	0.94	0.98	0.90	0.99	0.95	-

14.4 Y 染色体特异 STR 基因座在法医鉴定中的应用

1. 男女混合斑检验

性犯罪案件中，罪犯的精液与受害者的细胞成分常呈混合斑迹。虽然可以利用差异裂解的办法使精子细胞与女性成分分离，但有时仍会发生在精子 DNA 提取液中残留有女性物质。还有部分案件材料是男女混合血液或唾液时，不可能用差异裂解法分离开男女性 DNA，而应用 Y 染色体特异 STR 基因座分析，可以很简单地确定男性的单倍型。Y 染色体特异的 STR 基因座多态性分析和常染色体 STR 分析不同，即使是较大量的女性 DNA 成分也不抑制 Y 特异的 STR 等位基因的扩增。例如在由男性和女性混合物中，检验男性 DYS19、DYS390、DYS389 Ⅰ/Ⅱ 基因座，在男性成分仅占 1/2 000 时，即在 800ng 的女性 DNA 中只有 400pg 的男性 DNA 时，也能确定出男性的等位基因；而对照的常染色体 STR 系统如 VWA 基因座，只有男性成分超过 1/20 ~ 1/10 以上的比例时，才能确定男性的型别。对 Y 染色体特异 STR 基因座的检验，为性犯罪案件中混有被害者女性 DNA 的精斑或其他男性血液/体液与女性血液/体液的混合斑物证检验，提供了分析方法。

2. 父子关系鉴定

Y - STR 多态性是父系遗传，因此在父权鉴定中具有实际意义。Y - STR 在父子关系鉴定中的平均排除概率和个人识别概率是相同的。因此，Y - STR 座位连锁的多态性与常染色体基因座相比较，具有较高的排除能力。Y - STR 的父系遗传特征在被指控的父亲不能受检验的亲权关系鉴定中，可能提供有用的相关信息，例如被指控的父亲已经去逝，可以考虑采用该家庭中同一父系的男性亲属血样代替他作检验，提供关于他的 Y 染色体分子结构的全部信息。如果男孩的 Y 染色体与被指控父亲的兄弟、叔侄的 Y 染色体 STR 型不相同，显然可以排除。

3. 不同男性个体混合物（斑）分析

如精液 - 精液、血液 - 血液、血液 - 唾液等不同男性个体的混合物，提取的 DNA 显然是多个体混合材料，不同个体常染色体基因座的 DNA 很难区分，此时检验 Y 染色体特异遗传标记可能有所帮助。例如轮奸案件中的精斑，检验 Y 染色体特异的 STR 基因座常常可以确定精液供者的人数。此时虽然不能从混合物分析中直接推断出各个强奸犯的单倍型，但可以在对混合物和嫌疑人的单倍型分析直接比对时，有望获得罪犯的线索。

在个体识别中，当采用多个 Y 染色体特异 STR 基因座进行检验时，联合识别力可能很高。但 Y - STR 基因座等位基因是单倍体，各基因座的遗传是连锁的；同一父系的男性，具有同样的单倍型，因此 Y - STR 多态性作个人识别有一定局限性。有实际意义的是 Y - STR 单倍型不一致时，可以作出否定的结论。

14.5 X 染色体 STR

X 和 Y 染色体一样，都是性染色体，男性为 XY，女性为 XX。男性的 X 染色体来自母亲，与父亲无关，并将 X 染色体传递给女儿。女性的二个 X 一个来自父亲，一个来自母亲。父亲的 X 只传递给女儿，传递给儿子是 Y；而母亲的 2 个 X 染色体中任 1 个均有可能或传递给儿子或传递给女儿。所以，X 染色体的遗传既不同于常染色体，也不同于 Y 染色

体。性染色体上等位基因的传递均表现为性连锁（sex - linked）遗传特征，而其中 X 染色体的等位基因传递，表现为交叉传递。X 染色体的 STR 基因座，正常男性个体只有一个等位基因一条带，女性有二个等位基因，如二个等位基因不相同，表现为二条带，如二个等位基因相同则表现为一条带。

目前已报道了数个 X - STR 基因座多态性，如 HumDXS6807（DXS6807）、HUMARA（AGC）、HUMHPRTB 等基因座。

14.6 X 染色体的 STR 基因座

14.6.1 HumDXS6807 基因座

HumDXS6807 基因座（CDB：G00 - 365 - 390）定位于 Xpter - Xp22.2，和 HumARA，HumHPRTB 基因座不连锁，与此两基因座的遗传距离分别为 87 cM 和 151cM。表 14 - 18 为群体数据。

表 14 - 18　德国人 HumDXS60807 基因座群体数据

等位基因	长度(bp)	样本数		频　率	纯合子女性	总等位基因比例	频　率
		男性	女性				
11	251	114	295	0.479 + 0.014	65	409	0.496 + 0.012
12	255	4	14	0.023 + 0.004	0	18	0.022 + 0.004
13	259	3	6	0.010 + 0.003	0	9	0.011 + 0.003
14	263	48	145	0.235 + 0.012	16	193	0.234 + 0.010
15	267	32	134	0.218 + 0.012	12	166	0.201 + 0.010
16	271	6	15	0.024 + 0.004	0	21	0.025 + 0.004
17	275	2	7	0.011 + 0.003	0	9	0.011 + 0.008
		∑:209	∑:616		∑:93	∑:825	

基因座结构：引物 1 - N149 - （GATA） - GTAATGA - （GATA）2 - GAC - （GAATA）n - TGT - 引物 2n = 11 ~ 17

1. 引物序列　5′ - GAG CAA TGA TCT CAT TTG CA - 3′

5′ - AAG TAA ACA TGT ATA GGA AAA AGC T - 3′

2. 反应体系　在 25μl 反应体系中，含有 10ngDNA，1u Taq DNA 聚合酶，0.25μmol/L 每种引物，150μmol/L dNTPs，50mmol/L KCl，10mmol/LTris - HCl（pH9.0），1.5mmol/L $MgCl_2$。

3. 循环参数　94℃3min，1 个循环，94℃30s，51℃ 30s，72 ℃ 75s，28 个循环，最后 72℃ 6min。

14.6.2 HUMARA（AGC）基因座

HUMARA（AGC）基因座是 androgen 受体，定位于人类染色体 Xcenq13or4，重复序列为 AGC，已检出 15 个等位基因，等位基因片段为 263 ~ 308bp。中国汉族人群基因频率分布见表 14 - 19。

表 14－19　HUMARA（AGC）基因座等位基因分布频率

等位基因	中国汉族
263	0.0087
269	0.03478
272	0.0087
275	0.0000
278	0.01739
281	0.0087
284	0.1217
287	0.1391
290	0.1217
293	0.1217
296	0.0608
299	0.07826
302	0.07826
305	0.06957
308	0.06087
杂合度	0.89
DP	0.9051

1. 引物序列　5′－TCC AGA ATC TGT TCC AGA CGT GC－3′

5′－TTA CGC GTG CTG TGC AGG TTG CTG TTC CTC AT－3′

2. PCR 反应体系　50μl 或 20μl 反应体系，含有 50mmol/L KCl，10mmol/L Tris－HCl（pH8.3），1.5mmol/L $MgCl_2$，0.001mg/mL 白明胶，62.5μmol/L dNTPs，0.24μmol/L 每种引物，Taq DNA 聚合酶 0.5u。

3. PCR 热循环参数　95℃ 45s，65℃ 30s，72℃ 30s，25 个循环后，72℃ 延伸 5min。

14.6.3 HUMHPRTB 基因座

HUMHPRT 基因座是人次黄嘌呤磷酸核糖转移酶基因（human hypoxanthune phosphoribosyltransferase gene），定位于人染色体 Xq26.1，核心序列为 AGAT，已发现 12 个等位基因（表 14－20），片段长度范围为 257～303bp。

表 14－20　HUMHPRT 基因座等位基因频率

等位基因	频率		
	匈牙利人	美国白人	美国黑人
7		0.005	0
8	0.005	0.000	0
9	0.020	0.028	0.023
10	0.005	0.051	0.018
10M	0.005		
11	0.132	0.136	0.106
12	0.353	0.341	0.349

续表

13	0.284	0.229	0.211
14	0.137	0.150	0.202
15	0.054	0.047	0.073
16	0.005	0.014	0.018
杂合度		0.7346	0.7073

1. 引物序列　5′-ATG CCA CAG ATA ATA CAC ATC CC-3′
　　5′-CTC TCC AGA ATA GTT AGA TGT AGG-3′

2. PCR 反应体系　15μl 反应体系，含有 10~50ng 模板 DNA，50mmol/L KCl，10mmol/L Tris-HCl（pH 8.3），1.5mmol/L $MgCl_2$，0.001mg/mL 白明胶，200μmol/L dNTPs，1μmol/L 每种引物，Taq DNA 聚合酶 0.5u。

3. PCR 热循环参数　95℃ 45s，55℃ 45s，72℃ 2min，30 个循环。

14.6.4 DXS6799、DXS6804、DXS7132 基因座

X-STR 座位 DXS6799、DXS6804 和 DXS7132 在汉族人群分别检出 6，5 和 5 个等位基因。基因频率分布见表 14-21。

表 14-21　汉族女性 DXS6799、DXS6804、DXS7132 基因座等位基因频率分布

DXS6799		DXS6804		DXS7132	
片段大小	基因频率	片段大小	基因频率	片段大小	基因频率
252	0.0096	173	0.545	283	0.2328
256	0.0768	177	0.2182	287	0.4569
260	0.2788	181	0.3091	291	0.1810
264	0.3461	185	0.2272	295	0.1207
268	0.2596	189	0.1909	299	0.0086
272	0.0288				

1. 引物序列

（1）DXS6799　5′- ATG AAT TCA GAA TTA TCC TCA TAC C -3′
　　5′-GAA CCA ACCTGC TTT TCT GA-3′

（2）DXS6804　5′- CCC AGA TAT TTT GAC CAC CA -3′
　　5′-GGC ATG TGGTTG CTA TAA CC -3′

（3）DXS7132　5′-TTG CAT TCC GTA AAT TGT CCT-3′
　　5′-GTT ATC CCG AAT CGT ATT TGA CT-3′

2. PCR 反应体系　在 15μl 反应体系中 5×PCR 反应缓冲液（$MgCl_2$7.5mmol/L，NH_4Cl 25mmol/L，Tris-HCl 50mmol/L pH8.0，KCl 250mmol/L）3μl，10×dNTP 1.25mmol/L 1.5μl，$MgCl_2$（25μmol/L）0.4μl，引物Ⅰ（10mmol/L）、引物Ⅱ（10mmol/L）各 1.5μl，DNA 模板 20ng，Taq DNA 聚合酶 1u。

3. 循环参数　94℃ 4min，然后 94℃ 40s，55℃（DXS6799、DXS6804、DXS7132）或

58℃ 50s，72℃ 60s，35 个循环。

14.7 X 染色体 STR 的法医学应用

X－STR 为性连锁遗传，以单倍体形式传给子代。因此，X－STR 在应用上增加了特点，尤其是单亲父女关系的亲权鉴定。父子亲权关系可以通过父系遗传的 Y－STR 基因座分析直接确定父权关系，母子或母女关系鉴定可以通过母系遗传的线粒体 DNA 序列分析确定亲权关系。对于父女关系的单亲，既不能应用父系遗传的 Y 染色体基因座，也不能应用母系遗传的线粒体 DNA 分析，但可以用 X－STR 基因座分析父女单亲关系，父亲的 X 染色体必遗传给其女儿。如果女儿的 X－STR 基因座的等位基因与被假定的父亲不同，可以直接排除其父女关系，如果相同，则不能排除其父女关系，父权指数算法要比常染色体基因座简单些，但排除率也低很多。三联体或母子或母女关系鉴定，可参照常染色体基因座，但要注意 X－STR 基因座不是独立的，表现为连锁关系，以单倍型频率计算为宜。

另一个应用是个体识别。除具有与常染色体 STR 功能外，对于男性除精液（斑）外的体液（斑痕）或组织样品，单一男性样品 DNA 只有一个 X 染色体 DNA，X－STR 基因座只有一个等位基因。如果出现二个或二个以上等位基因则表明被检验样品混有其他个体 DNA，为混合样品。即使是同胞个体间 DNA 混合样品，也有可能被检出。

由于 X 染色体 STR 基因座多态性特征为 X 染色体伴性遗传，因此在计算评定 X 染色体 STR 基因座的法医遗传学指标 PE 值、Dp 值时，与常染色体的 STR 基因座的方法是不相同的。计算方法为：

当样品来自一个女性时：$DP = 1 - 2(\sum p_i^2)^2 + \sum p_i^4$

当样品来自一个男性时：$DP = 1 - \sum p_i^2$

进行亲权鉴定当孩子是一个女孩时，其 PE 值计算方法为：

$PE = 1 - \sum p_i^2 + \sum p_i^4 - (\sum p_i^2)^2$

当为男孩时：

$PE = 1 - 2(\sum p_i^2) + \sum p_i^3$

第十五章　线粒体 DNA（测序）分析

线粒体是存在于细胞质的一个含有双层膜的重要细胞器，是细胞的氧化中心和动力站。几乎人体所有细胞中均有线粒体，但不同组织细胞中线粒体的数目有所差异，如肝脏、心肌、骨骼肌等组织细胞中线粒体数目含量较多，每一个肝细胞中约有 1 000 ~ 2 000 个线粒体，可能与这些组织中的代谢率高有关。线粒体有自己的遗传系统，线粒体 DNA（mitochondrial DNA，mtDNA）是人类第二套基因组 DNA，也是人细胞中除核之外唯一含有 DNA 的细胞器。线粒体也有自己的蛋白质翻译系统和遗传密码。线粒体的基因组编码了 tRNA，rRNA 及一些功能蛋白质，如电子传递链酶复合体中的亚基，这些均参与维持线粒体系统的功能。由于受精卵中的线粒体几乎全来自于卵细胞（因为精子细胞中含有极少量的线粒体，只有 50 个拷贝），所以线粒体遗传系统往往表现为母系遗传，母亲所携带的 mtDNA 突变可传给她的子女。

15.1 人类线粒体基因组

mtDNA 是裸露的，不与组蛋白结合，存在于线粒体的基质内或粘附于线粒体内膜。在一个线粒体内往往有一至数个 mtDNA 分子。mtDNA 的自我复制也是以半保留复制方式进行的。

线粒体基因组是人类基因组的重要组成部分，剑桥大学的 Anderson 等人于 1980 年完成了对线粒体基因组的全序列测定，该序列可在基因库中找到，称之为 Anderson 序列或剑桥序列。现都以 Anderson 序列作为标准，进行人线粒体 DNA 序列多态性比对。现代的线粒体是从古代的一个细菌祖先进化衍生而来。人类线粒体基因组的序列共含 16569 个碱基对（其中 5122 A，5180 C，2171 G，4096T），为一条双链环状的 DNA 分子。双链中有一条为重链（H），即富含嘌呤链；另一条为轻链（L），即富含嘧啶链。这是根据它们的转录产物在 CsCl 中密度的不同而区分的。重链和轻链上的编码物各不相同（图 15 - 1），线粒体基因组共编码了 37 个基因，其中 22 种基因编码 tRNA 分子（用于线粒体 mRNA 的翻译），2 种编码 rRNA 分子（用于构成线粒体的核糖体 12S 和 16SrRNAs），另外还有 13 个编码氧化磷酸化电子传递链和 ATP 产生涉及的蛋白质和酶，它们都以 ATG（甲硫氨酸）的起始密码为开始，并有终止密码结构，长度均超过可编码 50 个氨基酸多肽的顺序。这 13 个可能的蛋白质均已被确定，其中 3 个为构成细胞色素 c 氧化酶（cytochrome oxidase，COX）复合体催化活性中心的亚单位，这三个亚基与细菌细胞色素 c 氧化酶相似，其序列在进化过程中高度保守，2 个为 F_0ATP 酶的 2 个亚基，7 个为 NADH - CoQ 还原酶复合体的亚基，还有一个是 CoH_2-细胞色素 C 还原酶复合体中的细胞色素 b 亚基。

图 15 - 1 表示了线粒体基因组两条链的编码情况，其中重链上的编码基因或编码产物有 12SrRNA、16SrRNA、ND（NADH - Cop oxidoreductase）1、ND2、ND4、ND4L、ND5、ND7 及 14 个 tRNA（图中用小写字母表示，表示其对应的氨基酸），还编码了 ATP 合成酶的第

6 亚基和第 8 亚基（图中表示为 A6、A8）。轻链编码了 ND6 及 8 个 *t*RNA。线粒体基因组与核基因组相比要经济或紧凑，核基因组中编码功能的序列不足 10% 而在线粒体基因组中只有很少非编码序列。在很多情况下，其 DNA 序列中几乎不含终止密码，在转录之后在 mRNA 转录物的特定区域加上多聚腺嘌呤核苷酸。

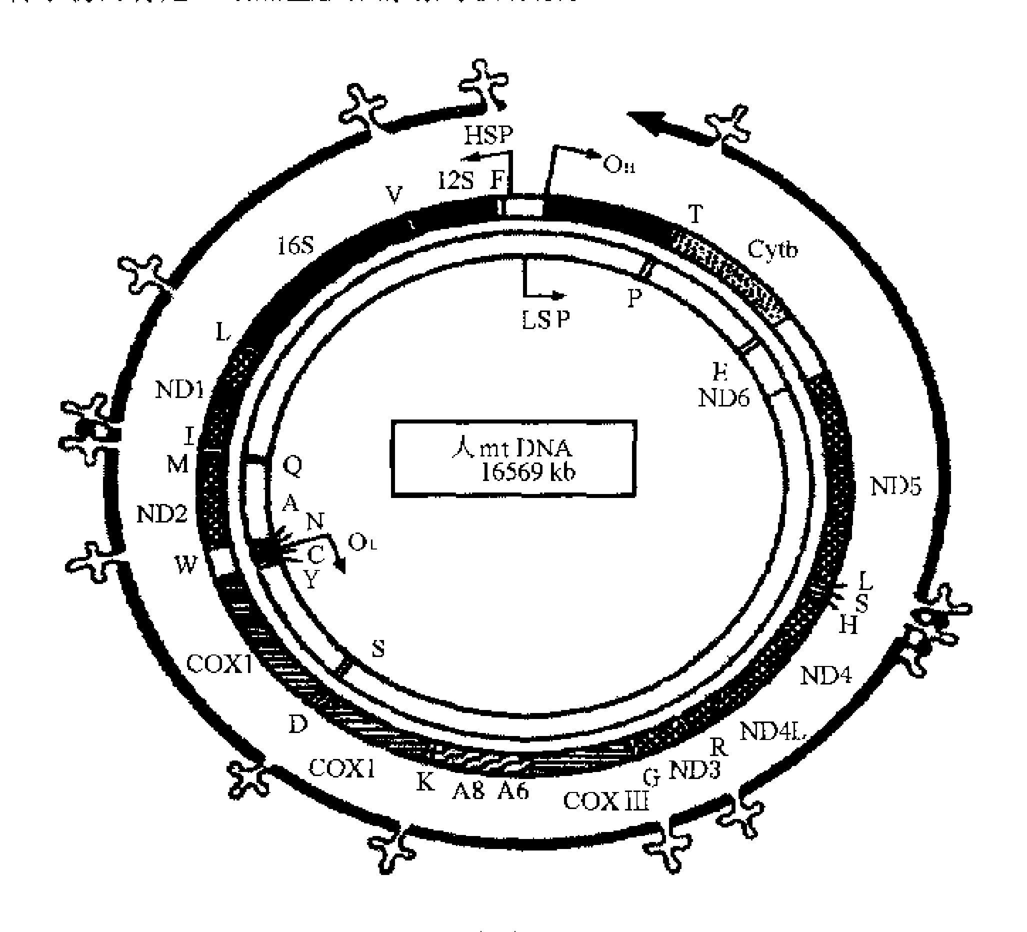


图 15－1 人线粒体 DNA 基因组编码图

最外圈为重链转录体；A：ATP 复合酶；COX：细胞色素氧化酶；
ND：NADH 氧化酶；HSP：重链转录启动子；LSP：轻链转录启动子

所有 mtDNA 的转录物都在线粒体内，所有 mtDNA 编码的蛋白质也在线粒体内，并在线粒体的核糖体上进行合成的。线粒体编码的 RNA 和蛋白质不被运出线粒体外，但构成线粒体核糖体的蛋白质是由胞质运入线粒体内的。用于蛋白质合成的所有 tRNA 也都是由 mtDNA 编码的。线粒体 mRNA 翻译的起始氨基酸为甲酰甲硫氨酸，这点与原核生物类似。

除编码区外，线粒体 DNA 还含有一个非编码区，叫控制区（control region，CR），又谓 D－环区（D－loop），这个名字得自于电子显微镜观察到的线粒体 DNA 在复制过程中形成的一种结构形状。控制区大约 1 100bp 或 1 125bp 大小，属于非编码区。在 D－loop 区中间，含有基因多顺反子 RNA 转录的 启动子（promoters）和重链复制的起点。Anderson 序列的碱基数字定位系统是从控制区中间附近开始的，因此控制区跨度为 nt16024－16519，然后nt1－576。相对于编码区，控制区在个体间存在较大差异，有许多单碱基取代与突变，mtDNA 控制区每 100bp 有 1～3 个核苷酸发生变化，另外还存在二个串联重复序列长度多

态性区，例如在第 514 - 523 位置的 CA 重复。碱基多态性主要集中在二个高变区（hypervariable region）：高变区Ⅰ和区Ⅱ，简称 HVRⅠ和 HVRⅡ（图 15 - 2）。高变区Ⅰ范围为 nt16024 ~ 16365，高变区Ⅱ为 nt73 ~ 340，高变区边界不是固定的，不同实验室或研究报道有所不同。1998 年 Lutu 等人报道线粒体 DNA 还存在高变区Ⅲ（nt438 ~ 574），二核苷酸 CA 重复位于这 HVRⅢ。HVRⅠ的多态性最好，重链的复制区起点在 HVRⅡ。mtDNA 法医学价值就在于 D - loop 区的个体间的序列差异，主要是 HVRⅠ和 HVRⅡ。

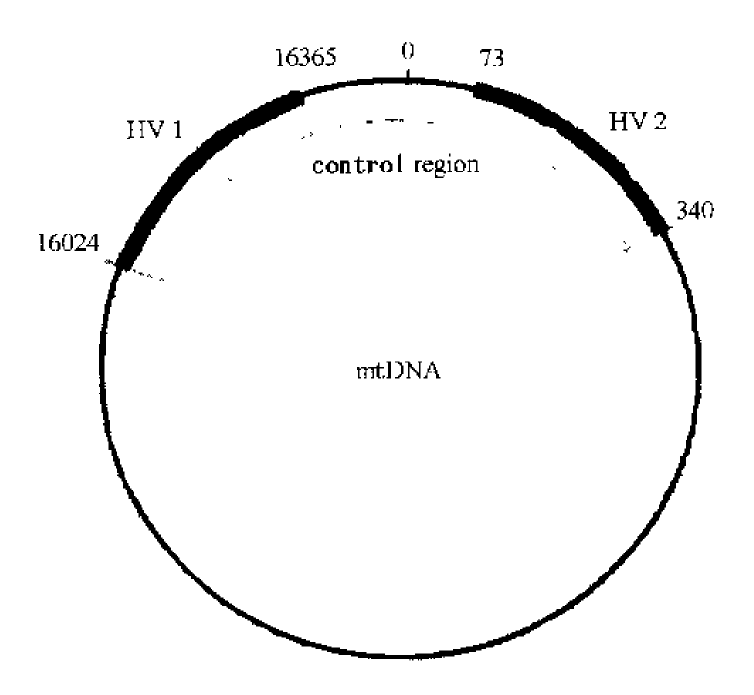


图 15 - 2　人线粒体 DNA 控制区的两个高变区

下列序列是 D - 环的碱基序列：

1 gatcacaggt ctatcaccct attaaccact cacgggagct ctccatgcat ttggtatttt

61 cgtctggggg gtatgcacgc gatagcattg cgagacgctg gagccggagc accctatgtc

121 gcagtatctg tctttgattc ctgcctcatc ctattattta tcgcacctac gttcaatatt

181 acaggcgaac atacttacta aagtgtgtta attaattaat gcttgtagga cataataata

241 acaattgaat gtctgcacag ccactttcca cacagacatc ataacaaaaa atttccacca

301 aacccccccct cccccgcttc tggccacagc acttaaacac atctctgcca aaccccaaaa

361 acaaagaacc ctaacaccag cctaaccaga tttcaaattt tatcttttgg cggtatgcac

421 ttttaacagt caccccccaa ctaacacatt attttcccct cccactccca tactactaat

481 ctcatcaata caacccccgc ccatcctacc cagcacacac acaccgctgc taaccccata

541 ccccgaacca accaaacccc aaagacaccc cccacagttt atgtagctta cctcctcaaa

601 gcaatacact gaaaatgttt agacgggctc acatcacccc ataaacaaat aggtttggtc

661 ctagcctttc tattagctct tagtaagatt acacatgcaa gcatccccgt tccagtgagt

721 tcaccctcta aatcaccacg atcaaaagga acaagcatca agcacgcagc aatgcagctc

781 aaaacgctta gcctagccac acccccacgg gaaacagcag tgattaacct ttagcaataa

841 acgaaagttt aactaagcta tactaacccc agggttggtc aatttcgtgc cagccaccgc
901 ggtcacacga ttaacccaag tcaatagaag ccggcgtaaa gagtgtttta gatcaccccc
961 tccccaataa agctaaaact cacctgagtt gtaaaaaact ccagttgaca caaaatagac
15961 gaaaaagtct ttaactccac cattagcacc caaagctaag attctaattt aaactattct
16021 ctgttctttc atggggaagc agatttgggt accacccaag tattgactca cccatcaaca
16081 accgctatgt atttcgtaca ttactgccag ccaccatgaa tattgtacgg taccataaat
16141 acttgaccac ctgtagtaca taaaaaccca atccacatca aaaccccctc cccatgctta
16201 caagcaagta cagcaatcaa ccctcaacta tcacacatca actgcaactc caaagccacc
16261 cctcacccac taggatacca acaaacctac ccacccttaa cagtacatag tacataaagc
16321 catttaccgt acatagcaca ttacagtcaa atcccttctc gtccccatgg atgacccccc
16381 tcagataggg gtcccttgac caccatcctc cgtgaaatca atatcccgca caagagtgct
16441 actctcctcg ctccgggccc ataacacttg ggggtagcta aagtgaactg tatccgacat
16501 ctggttccta cttcagggtc ataaagccta aatagcccac acgttcccct taaataagac
16561 atcacgatg

15.2 线粒体 DNA 遗传系统的特点

mtDNA 的遗传系统表现出与核 DNA 遗传系统不完全相同的特点。

1. 多拷贝　在正常细胞中，一个细胞的细胞核中只有一套染色体或核基因组，一个基因座位上有两个等位基因，一个来自母亲，另一个来自父亲。然而，每个细胞内则含有多个线粒体，并且每个线粒体内含多个拷贝的线粒体基因组 DNA（血小板和未受精的卵子例外，它们中的每个线粒体内只含有一个拷贝的 mtDNA），体细胞大约含有 200 ~ 1700 个 mtDNA 拷贝。组织类型不同，数目也不同。由于有相对丰富的 mtDNA，因此在核 DNA 已降解的样品中仍有较大可能地提取到 mtDNA。

2. 高突变率　mtDNA 具有比核 DNA 更高的突变率，mtDNA 中某些区域的进化速度是核基因组的 6 ~ 17 倍。mtDNA 的非编码区在个体间表现出高度的差异性，例如，高加索人在 mtDNA 的高变区 Ⅰ 和 Ⅱ 中平均有 8 个核苷酸不同。

3. 异质性　一个个体的 mtDNA 有两种或两种以上不同碱基序列差异的情况称之为异质性（heteroplasmy）或叫杂合性；一个个体的所有 mtDNA 分子都相同，则称之为同质性（homoplasmy）或 同一性。个体 mtDNA 异质性的存在形式有三种：（1）个体的同一组织中含有一种以上的 mtDNA 序列；（2）同一个体不同组织的 mtDNA 序列不同，一种组织含有一种 mtDNA 序列，而在另一组织中存在着另一种 mtDNA 序列；（3）个体的一种组织中的 mtDNA 为同质性，而在另一种组织表现为异质性。在细胞发育过程中，mtDNA 分子彼此独立复制，不涉及细胞减数分裂，因此本质上没有重组。而且 mtDNA 复制过程中存在比核基因组 DNA 更高的出错率。这些因素导致个体中 mtDNA 分子存在着差异，并且不同的变异体独立复制、分配。但是 mtDNA 传递过程中的遗传瓶颈限制了 mtDNA 变异体在后代中的传递，使其在人群中分布没有那么广泛，只有 2% ~ 8% 人群存在异质性。控制区存在异质性首先是在沙皇尼古拉 Ⅱ 世骨骸鉴定中发现的，在 16169 核苷酸位置尼古拉 Ⅱ 世骨骸存在两种碱基 C 和 T。异质性经常出现在 16093、16129、16169（C/T）、16311（A/G）、

16293（C/T）、16355（C/T）、16397、00309 等位置。异质性出现的二个热点：一个是 16093 位置，频率 34%，另一个是 16129 位置，频率为 17%，其他位置的异质性只观察到一次。另外二个多聚 C－stretch 中还存在长度异质性，一个在 HVRⅠ，一个在 HVRⅡ。在 HVRⅠ，大多数在 16189 位置有一个碱基 T，隔断了左右两边的一串碱基 C，如果 16189 位置碱基 T 被 C 取代，形成一连续的 C，使人群中个体之间存在 C－stretch 长度差异。在 HVRⅡ，在 303～315 碱基位置也存在 C－stretch 的长度异质性，这是Ⅱ区的重要差异性。异质性的存在虽然使法医学鉴定复杂化，但有其鉴定意义。在实际应用中必须考虑异质性的存在，了解异质性出现的位置，有助于指导结果解释。由于调查的有异质性的样品数太少，尚未获得精确的大多数核苷酸位置异质性发生的频率。

异质性分析方法有变性高压液相色谱（denaturing highperformance liquid chromatography，DHPLC），变性梯度凝胶电泳（denaturing gradient gel electrophoresis，DGGE）、单链构象多态性（single－strand conformation polymorphism，SSCP）。DGGE 的灵敏度或分辨率比直接测序高，分析 253 个体序列共检出 35 个个体的异质性，而用直接测序方法只检出 3 个个体的异质性。

mtDNA 的高突变率及核苷酸的取代是异质性产生的原因，不同位置的碱基取代速度明显不同。在 mtDNA 从亲代向子代传递过程中涉及一个或多个遗传瓶颈，造成同一个体不同组织中异质性比例不同、同一母系成员中 mtDNA 不同，甚至变成一个同质性的核苷酸取代，同一母系家属 CR 序列可能相互不同，Mark（1997）报道有一个家系，母亲和女儿在 16355 位置存在 C/T 异质性，血液 mtDNA 中主要是 C，伴有相对弱的 T，在唾液细胞 mtDNA 也发现 16355C/T 异质性，C、T 强度比较平衡；分析其毛发，发现一根毛发中为 16355C，第二根为 T，而第三根为 C/T 共存，C 和 T 峰强度相当。用同样方法分析其儿子 mtDNA，结果发现儿子的毛发和唾液 mtDNA 序列在 16355 位置均为 T。因此，用 mtDNA 分析方法进行母系成员鉴定对时，要考虑到异质性的存在，以避免假排除。

4. 母系遗传　在精卵结合的受精过程中，受精卵中几乎所有的 mtDNA 来自于卵子，表现为母系遗传（maternal inheritance）。主要原因是简单的数量问题，因为精子头部不含或极少有 mtDNA，而卵子则含有上百万个拷贝 mtDNA，另外有学者认为还存在一个特异的识别机制，减少了进入到卵子的父方 mtDNA。有人实验将来自精子细胞的 mtDNA 引入到没有 mtDNA 的培养的体细胞中，含有功能的 10%～20% 精子线粒体迅速进入到培养细胞中，但只有极少数细胞（$1/10^5$）存活 48h 以上，而来自自体细胞的线粒体引入到培养细胞中，很快取代了外源 mtDNA，这证明存在着一种精子线粒体减少的机制。因此，mtDNA 分子的传递与孟德尔遗传不同，表现为母系遗传。母亲将携带的突变 mtDNA 分子传给她所有的子女，再由她的女儿能把这种突变传给下代，四代内没有变异发生。

5. 不同的遗传密码　线粒体的遗传密码与核的遗传密码不完全等同。

因此，mtDNA 与核 DNA 具有较大差异，表 15－1 为两者的比较。

表 15-1 人核 DNA 和 mtDNA 的比较

特点	核 DNA	mtDNA
大小	~3 百万 bp	16569bp
每个细胞的拷贝数	2	多于 1 000
结构	线形；包装在染色体中	环状
遗传来源	父母两方	母亲遗传
生殖重组	有重组	没有重组
突变率	低	至少是核 DNA 的 5~10 倍
全序列获得	人类基因组工程完成	1981 年

15.3 mtDNA 遗传多态性分析方法概述

15.3.1 分析方法

mtDNA 的多态性主要表现为 D-loop 区域的碱基序列的差异性，除此还有另外 2 个串联重复单位的长度多态性。有多种方法检测 D-loop 区多态性，每个方法各有特点，表 15-2 介绍了各种方法。由于单个细胞中含有数 10 万个拷贝线粒体 DNA，使之比核 DNA 更容易进行测序，且 mtDNA 单倍体的性质使之比二倍体核 DNA 更容易测序，加上直接测序可以获得更多的信息，因此目前最为常用的是 mtDNA 的直接测序。

表 15-2 检测 mtDNA 差异的方法

方法、技术	简介
序列特异寡核苷酸（SSO）（点杂交分析）	23 个 SSO 探针检测 mtDNA 控制区的 9 个区域，在 525 个体中发现 274 种 mtDNA 分型；17 个 SSO 探针反向杂交检测 mtDNA HVRⅡ的多态性。可在 PCR 扩增后 2.5h 内获得序列结果，方法简便、快速。
固相微型测序（solid-phaseminisequencing）	检测控制区内 10 个取代多态性和 2 个长度多态性
单链构象多态性（SSCP）	在非变性凝胶上检测 DNA 单链构象多态性
低特异单个特异引物 PCR（LSSP-PCR）	正常 PCR 后，单个引物和低退火温度下扩增出具有个体专一的 PCR 图谱；分析 30 个个体，它们在控制区的所有签名图谱不同。
PCR-限制片段长度多态性（PCR-RFLP）	HVR Ⅰ 的 199bp 区域用 RsaI 消化
变性梯度凝胶电泳（DGGE）	在变性梯度凝胶上电泳，杂合二倍体在凝胶中移动较慢，可以与纯合二倍体分开；可以区分单一位置不同的二个个体。
高密度 DNA 阵列（DNA chip or microarray）	13 500 探针完全与整个 mtDNA 基因互补，固定在一个微阵列上，通过杂交平行检测。
直接测序	利用 PCR 制备测序模板，对 mtDNA 的高变区Ⅰ和Ⅱ进行手动或全自动测序，获得整个区域所有核苷酸序列。

15.3.2 遗传统计分析

mtDNA D-loop 区域的碱基序列为连锁遗传，只能算作一个基因座，以单倍型表示，进行遗传群体分析按照单倍型频率统计计算。

基因差异度（Gene diversity，GD）来表示 mtDNA 的个体识别能力。计算公式为：

$GD = [n(1-\sum x^2)]/(n-1)$

式中，n 为群体的样品数式中，x 为线粒体 DNA 单倍型在群体中的频率。

15.4 线粒体 DNA 测序

线粒体 DNA 测序采用的是 Sanger 双脱氧终止法，随着荧光标记自动测序技术的发展，以及 PCR 技术引入到测序中，目前法医 DNA 分析较为常用的线粒体测序方法是：利用 PCR 技术扩增制备测序模板，采用循环测序荧光标记自动检测测序。荧光标记自动测序方法有引物标记（dye primer）和末端标记的（dye terminator）两种方法。dye primer 利用 M13 等通用引物进行测序反应，使测序的每一片段均带有引物染料，dye terminator 利用荧光标记的双脱氧核苷酸在测序反应中掺入到各测序片段。PE 公司开发的 BIG DYE 试剂盒，每个染料分子带有一个放大因子，测序反应灵敏度比以前任何测序试剂盒的高，目前普遍采用此试剂盒进行测序。

15.4.1 测序引物

mtDNA 多态性主要集中在 HVR Ⅰ 和 Ⅱ，在这二个区内存在相当多的保守区可以设计引物，表 15－3 是报道的部分线粒体 DNA 高变区 PCR 模板制备和测序的引物对。

表 15－3　引物与产物

区域	引物对	产物大小（bp）
HVR Ⅰ	L15929/H16498	608
	L15997/H16401	405
	F15971，15997/R16410	440，433
	F16159/R16397	239
	F15971/R16255，R16258	285，288
	F16144/R16410	267
HVR Ⅱ	F15/R381，R389，R484	367，375，470
	F29/R381，R389，R408，R484	353，361，380，456
	F15/R270，R274，R285，R389	256，260，271，275
	F29/R270，R274，R285	242，246，257
	F155/R381，R389，R484	227，235，330
引物序列		
L15929	5′CAC CAG TCT TGT AAA CCG GA3′	
F15971	5′TTA ACT CCA CCA TTA GCA CC 3′	
L15997	5′CAC CAT TAG CAC CCA AAA GCT 3′	
R16255	5′CTT TGG AGT TGC AGT TGA TG 3′	
R16258	5′TGG CTT TGG AGT TGC AGT TG 3′	
F16144	5′TGA CCA CCT GTA GTA CAT AA 3′	
16401	5′TGA TTT CAC GGA GGA TGG TG3′	

续表

R16410	5'GAG GAT GGT GGT CAA GGG AC 3'
H16412	5'GTG CGG GAT ATT GAT TTC AC 3'
H16498	5'CCT GAA GTA GGA ACC AGA TG 3'
L00029	5'GGT CTA TCA CCC TAT TAA CCA C3'
F15	5'CAC CCT ATT AAC CAC TCA CG 3'
R270	5'TGG AAA GTG GCT GTG CAG AC 3'
R274	5'TGT GTG GAA AGT GGC TGT GC3'
R285	5'GTT ATG ATG TCT GTG TGG AA 3'
F155	5'TAT TTA TCG CAC CTA CGT TC 3'
R381	5'GCT GGT GTT AGG GTT CTT TG 3'
R389	5'CTG GTT AGG CTG GTG TTA GG 3'
H00408	5'CTG TTA AAA GTG CAT ACC GCC A3'
R484	5'TGA GAT TAG TAG TAT GGG AG 3'
R00484	5'TGA GAT TAG TAG TAT GGG AG 3'

引物序列位置根据 Anderson 序列确定，F 代表正链即线粒体 DNA 轻链；R 为反向链（线粒体重链）

测序时可以用 PCR 反应中正向引物或反向引物，也可以使用 PCR 产物内部序列互补的引物。

15.4.2 利用 BIG DYE 试剂盒进行测序

BIG DYE 试剂盒利用的是末端染料标记的方法，荧光染料分子标记在 ddNTP 上，在测序反应中，带荧光染料的 ddNTP 分子作为底物被掺入到每一测序片段，使每一片段均带有荧光标记。由于荧光标记在反应过程中是由反应底物带入的，因此对测序反应中的引物没有特殊要求，不需要连接 M13 的通用序列。

1. 各种检材 mtDNA 的提取

采用 Chelex－100 方法提取 mtDNA，详见 DNA 提取章节。

2. mtDNA D－loop 区的目的片段 PCR 扩增

在 50μl 反应体系中，含有 10μmol/L Tris－HCl（pH8.3），50μmol/L KCl，1.5μmol/L $MgCl_2$，200μmol/L dNTPs，1μmol/L 每种引物，1.25u Taq DNA 聚合酶，10μl Chelex－100 提取液。

PCR 热循环参数：94℃ 45s，50℃ 1min，72℃ 3min，循环 32 次，最后 72 ℃延伸 6min。

PCR 扩增产物检测：取 5μl 扩增产物加在 5％聚丙烯酰胺凝胶中，200v 电压下电泳 30min，银染法显色，观察是否得到扩增产物，是否具有特异性。

对于毛发和骨骼 mtDNA 模板，由于量少，为了得到足够产量的测序模板，常用 nested－PCR，以提高成功率。

Nested－PCR 的体系，包括第一步扩增和第二步扩增：

第一步扩增：25μl 反应体系中，含有 Chelex－100 提取的模板 DNA 5μl，200μmol/L dNTPs，0.5μmol/L 引物 L15997 和 H16401，10×缓冲液 2.5μl，1u Taq DNA 聚合酶。然后放入扩增仪中，进行扩增。循环参数为：94℃ 1min，56℃ 30s，72℃ 2min，循环 20 次后，72℃保温 7min，最后 4℃。如果使用血液或血痕提取的模板 DNA，那么只需 2μl。

第二步扩增：在 50μl 反应体系中，取第一步扩增产物 DNA 1～2μl，200μmol/L dNTPs，0.5μmol/L 引物 L16072 和 H16328，10×缓冲液 2.5μl，1u Taq DNA 聚合酶。然后放入扩增仪中，进行扩增。循环参数为：94℃1min，56℃30s，72℃90s，循环 30 次后，72℃保温 7min，最后 4℃。

3.mtDNA 目的片段的纯化

确认获得了 mtDNA 目的片段的扩增产物后，需将 mtDNA 目的片段从 PCR 扩增产物分离出来，与 PCR 反应中的引物等分开，因为 PCR 反应引物干扰测序反应。纯化的方法很多，目前较常用的是：（1）采用的 Promega 公司的 Winzard 试剂盒纯化；（2）Microcon 100 超滤法（3）QIAGene 试剂盒纯化。

（1）Promega 公司的 Winzard 试剂盒纯化法

该方法利用的是试剂盒中的硅胶具有吸附核苷酸能力及其在不同有机溶剂下对核苷酸结合率不同的特点，藉以纯化目的片段。方法如下：

①将 PCR 产物全部（30～50μl）加到 1.5ml 的离心管中，加入 100μl 试剂盒所带的纯化缓冲液，混匀。

②加 1ml 纯化树脂（硅胶），混匀 1min；

③用一次性 5ml 注射器针管套在过滤柱上，然后将混合液加到注射器中，缓缓向下推活塞，保持流速为 2 滴/s，使混合液通过过滤柱，DNA 片段被树脂结合；

④拔下过滤柱，抽出活塞，再将注射器针管套在过滤柱，加入 2ml 80%异丙醇，缓缓向下推活塞，保持流速为 2 滴/s，使异丙醇通过过滤柱，冲洗过滤柱；

⑤拔下过滤柱，将过滤柱套在 1.5ml 离心管中，100 00r/min 离心 2min，以去除异丙醇；

⑥拔下过滤柱，将过滤柱套在另一干净的 1.5ml 离心管中，加 30～50μl TE 或超纯水于过滤柱中，于室温放置 2min 后，10 000r/min 离心 20s，离心后的流出液作为测序反应模板。流出液中不含有引物。

（2）Microcon 100 超滤法

其原理是利用超滤膜的不同孔径的微孔允许比孔径小的分子通过膜，流到超滤膜下面，而大于孔径的分子仍保留在超滤膜上面，使大、小分子分离开来。Microcon100 截留分子量为 100 000，引物等小分子通过滤膜，目的片段（大于 100 000DLT）保留在膜上，反向离心可以收集目的片段。方法如下：

①将 Microcon 100 套在 1.5ml 离心管中（产品随带），将 PCR 扩增产物全部加在超滤管中；

②2 000r/min 离心 15min；

③去掉 1.5ml 离心管，将超滤管反套在另一干净的 1.5ml 离心管（又叫收集管）；

④4 000r/min 离心 2min，收集液体，可直接用于测序反应。

硅胶吸附法回收率及纯度没有 Microcon 100 超滤法高，但前者成本比后者底。注意事项请参考 DNA 纯化部分。

4. 测序反应

在 10μl 反应体系中进行测序反应：

取 4μl master mixture（包含 4 种 dNTP，4 种 ddNTP，缓冲液，FS 测序级酶），引物（一条）3.2pmol，30～90ng 纯化好的 PCR 产物模板 DNA（一般为 1～2μl），加超纯水至总体积为 10μl。放入扩增仪中进行测序反应。循环参数为：94℃1min，50℃5s，72℃4min，循环 25 次。

5. 测序反应产物的纯化

测序反应结束后，在反应液中加入 90μl 66.7%异丙醇沉淀 DNA，充分混匀 5～10min，13 000r/min 离心 20min，收集沉淀，再用 250μl 66.7%异丙醇洗沉淀，13 000r/min 离心 20min。去上清液，沉淀在室温下放置 30min，使异丙醇挥发掉，溶于 4μl 上样缓冲液中，取 1μl 进行测序。

6. 测序胶的制备（以下步骤均以 ABD XL377DNA 自动测序仪为例）

自动测序的优点之一是荧光发色基团吸收了激光束提供的能量而发射光信号，信号被探测器检测并传送至计算机，经整理显示出四种彩色图像。其缺点是凝胶本身及胶板上的任何污染物所放射出的荧光均会被当作信号收集。所以，在序列胶制备的过程中，很关键的是要消除一切可能产生荧光的物质。

（1）胶板的清洗和组装

洗制胶玻璃板时可用塑料刷（公司专门配备）或松软的海绵沾少量洗涤剂擦洗，用水反复多次冲洗，以除去残留的洗涤剂，再用蒸馏水冲洗干净，空气中晾干，晾干后玻璃板灌胶面肉眼应看不到任何痕迹。用 0.4mm 的夹条（spacer）装配好玻璃板。两边用夹子对称夹好，调整制胶玻璃板水平后，即可灌胶。

（2）6%聚丙烯酰胺－尿素胶的配制

核苷酸序列的数量和质量都取决于聚丙烯酰胺的质量，而聚丙烯酰胺浓度则取决于待分析 DNA 片段的长短。6%聚丙烯酰胺胶可读出距离引物 5′端 25～400bp，如果要读出距引物 5′端更远的序列，可选用 4%～5%的聚丙烯酰胺来测定。

①试剂

10×TBE：194g Tris（0.8mol/L），99g 硼酸（0.8mol/L），7.44g EDTA（0.01mol/L），加去离子水至 200ml

40%聚丙烯酰胺：190g 丙烯酰胺（测序级），10g N，N′－甲叉丙烯酰胺，加去离子水 500ml，搅拌 20min，加一汤匙 MB－树脂（可以不加），磁力搅拌 30min，过滤，装入棕色瓶内，置 4℃保存。也可用商品化的 50% Long Ranger 胶溶液，直接使用。

10%过硫酸铵（W/V）：1.0g 过硫酸铵加超纯水 10ml，溶解后置 4℃。

TEMED

尿素

②含 6mol/L 尿素的 6%聚丙烯酰胺凝胶配制（30ml）

称取 10.8g 尿素，加 15ml 超纯水，稍加热，完全溶解。加 3ml 10×TBE，3.6ml 50% Long Ranger 胶溶液，混匀，用 0.45μm 孔径的过滤器过滤，去除杂质。再加入 150μl 10%过硫酸铵，15μl TEMED。混匀。

用 50ml 注射器吸入胶液，从玻璃板一侧将之灌注到两玻璃板之间的孔隙中。连续缓慢灌注，避免出现气泡。待胶液注满玻璃板后，胶板的上方插入鲨鱼齿梳子，平端靠胶

液。室温下放置 1～2h，固化聚合。灌注胶也可用平推法，详见 STR 章节。

制备凝胶板时应注意：（1）在配制胶液和制胶时，要带手套。因为丙烯酰胺有强烈神经毒素，经皮肤吸收。在称取粉末时，要注意防止从呼吸道吸入。（2）如果用乙醇擦洗玻璃板时，一定要等乙醇挥发尽，才可制胶。（3）灌胶用的注射器，每次用后，用热水冲洗，再用蒸馏水冲净，以防堵住。

7. 电泳

用纸擦去聚合好的凝胶玻璃板外面的脏物，清洁玻璃板，将清洁好的凝胶板固定在仪器上，开机，启动 ABI PRISM 377XL Collection，用 CHECK A 模式检查玻璃板及凝胶是否干净，标准是扫描基线低且平整，如果基线不平整，取下凝胶板，重新擦洗扫描口附近的两面玻璃板，再检查，直至达到要求，有时是凝胶中的脏物产生的荧光扫描线，不论如何擦洗，仍然去除不尽，此时可以不管它，进行预电泳，这些脏物可以电泳除去。

在上下电极缓冲液槽内加入 1×TBE 缓冲液，夹好保温板，设置各种参数，进行预电泳。

预电泳模式（Prerun modual）：Seq PR36A－2 400

电压：1 000V

温度：51℃

2400scans/hr

Virtual set A

电泳模式（Run modual）：Seq Run 36E－2 400

电压：3 000V

温度：51℃

2400scans/hr

Virtual set E

预电泳 10～15min。此时准备上样样品，将样品在 95℃变性 2min，迅速放入冰水浴中。

取 1μl 变性样品加在凝胶中，开始电泳，收集数据，在 3 000V 下电泳 3h。

8. 数据分析

打开凝胶文件，按加样顺序给各泳道划线，抽提各泳道凝胶数据，正确选用 Matrix 文件，然后分析，读取结果（图 15－3）。

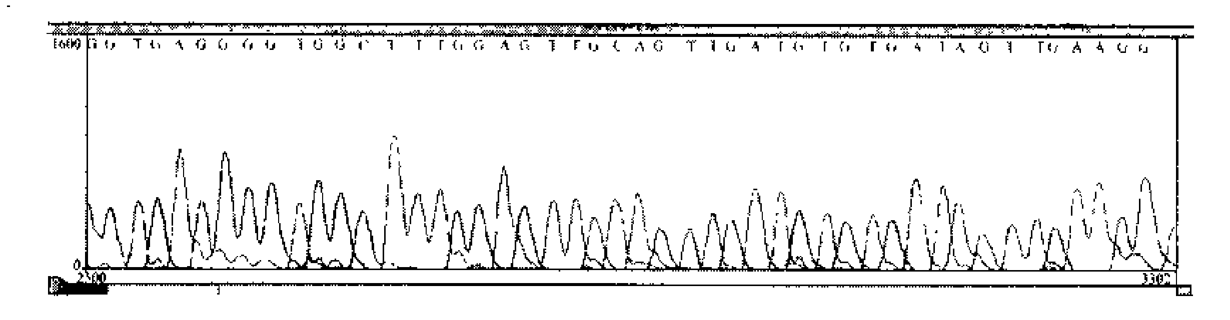


图 15－3　测序结果

9. 差异性比对

运用 PE 公司 Sequence Editor 软件，将所检测样品 DNA 的碱基序列与 Anderson（1981）的序列比对，可以得到该样品线粒体 DNA 高变区碱基序列在确定位置的差异，从而得到不同个体间线粒体 DNA 高变区碱基序列差异。结合电泳图谱，复核结果（图 15－4）。

图 15－4 待测样品与标准序列的比较结果

图中“*”表示与标准序列不同，“－”表示与标准序列相同

10. 分型

以第一个报道的人 mtDNA 序列（Anderson 序列）为标准，进行 mtDNA 分型。当一个个体的序列与 Anderson 序列比较时，只记录与标准序列不同的位置和核苷酸。例如 HVR Ⅱ中的第 73 位置核苷酸，Anderson 的序列为 A，测序结果该样品为 G，那么这个人的序列表达为 73G，如果没有碱基或位置被注明，那么显然表示这个序列与 Anderson 的序列相同；如果某一位置的碱基不清楚，那么在这个碱基位置数后标记‘N’，例如 16125N，表示在 16125 处的核苷酸序列不能确定；对第一个插入，表达为在插入碱基后加一个小数点和数字即‘.1’，对第二个插入则为小数点加 2 即‘.2’，依次类推；对于单一多聚体特点，不了解发生的插入的真正位置的情况下，那么假定插入发生在单一多聚体的最后位置。例如单一多聚体区域，插入较多发生在 311 和 315（包括 311、315）之间，多态性的 C 插入被认为在 315 后，因此对这个插入命名为 315.1C；同样的，对于缺失，在表示缺失位置后加‘d’，例如 220d，表示在 220 处核苷酸发生了缺失。对于确定的二异质体，此位置碱基按主要成分碱基记录。如果只怀疑但不确定有异质体，表达为‘N’，例如 A/C 的异质体可以表示为‘N’或‘R’。

15.4.3 采用通用引物 M13 荧光标记法测序 HV1 区域

采用 M13 引物荧光标记系统检测，要求模板制备的 PCR 反应引物 5′带有 M13 序列，在 PCR 反应中产物与 M13 互补的序列，作为测序反应中引物退火结合位置。

1. 引物序列

引物 1：M13（－21）L15997：5′－TGTAAAACGACGACGGCCAGTCACCATTAG CACCCAAAGCT－3′

划线部分为 M13（－21）通用序列，针对应用 PE－ABD 公司的引物荧光标记测序试剂盒而设计的，不影响 PCR 反应。

引物 2：H16401：5′－TGATTTCACGGAGGATGGTG－3′

2. 测序模板纯化及后面的测序反应产物纯化方法同前面介绍的 BIGDYE 方法。

3. mtDNA 目的片段 DNA 序列测定——荧光标记的双脱氧终止法

（1）标记反应管：每一个样品平行设置标记 A、T、G、C 的 4 个 PCR 反应管；

（2）按下列准备测序反应：

	A 反应管	T 反应管	G 反应管	C 反应管
反应体积（μl）	5	10	10	5
纯化好的待测 DNA	1μl	1μl	1μl	1μl
5×测序缓冲液	1μl	2μl	2μl	1μl
Taq DNA 聚合酶	2.5u	2.5u	2.5u	2.5u
荧光标记的 M13（－21）引物	0.8pmol/L JOE 标记的腺嘌呤特异的 M13（－21）引物	1.6pmol/L ROX 标记的胸腺嘧啶特异的 M13（－21）引物	1.6pmol/L TAMRA 标记的鸟嘌呤特异的 M13（－21）引物	0.8pmol/L FAM 标记的胞嘧啶特异的 M13（－21）引物
2′－双脱氧核苷酸－5′－三磷酸混合液	1μl 2′－双脱氧腺苷酸	1μl 2′－双脱氧胸腺苷酸	1μl 2′－双脱氧鸟苷酸	1μl 2′－双脱氧胞苷酸

（3）混匀上述反应液。

（4）在热循环仪中进行测序反应。95℃ 1min，50℃ 30s，72℃ 30s，循环 15 次；95℃ 1min，72℃ 30 s，循环 15 次，迅速进入 4℃。

（5）合并同一样品的 A、T、G、C 四个反应管产物。

（6）加 80μl －20℃无水乙醇和 1.5μl 3mol/L NaCA，沉淀 DNA 15min。

（7）16 000r/min 离心 20min，收集沉淀。

（8）干燥沉淀物，用按 5∶1 混合的去离子甲酰胺∶50mmol/L EDTA－Na_2（pH8.0）溶解。

（9）95℃变性 2min，迅速放置于冰水浴中。

（10）用 PE－ABD 373A，377 测序仪测序，样品加在 6%聚丙烯酰胺凝胶中，或用 ABD310 毛细管测序仪测。若用毛细管测序仪，选用 POP6。

（11）测序结果用测序软件自动分析，步骤同末端标记法。

15.4.5 循环测序两种荧光标记法测序的峰形特点

由于染料掺入方式不同，两种荧光标记法测序图中碱基的信号强度有差异，峰形特点不同。了解各自的特点，有助于序列的阅读。

1. 引物标记法测序的峰形

（1）在 2 个或多个连续 C 中，第一个 C 的峰一般比较低，尤其是后面的碱基为 G 时。

（2）在 2 个或多个连续 A 中，后面紧跟碱基 C 的第一个 A 的峰较大。

（3）在 2 个或多个连续 C 中，第二个 C 峰的信号被放大。

2. 末端标记法测序的峰形

（1）碱基 G 后面 C 峰一般较弱（图 15－5），当碱基 G 后面紧跟一串碱基 C 时，第一个 C 的这种情况更加明显（图 15－6）。

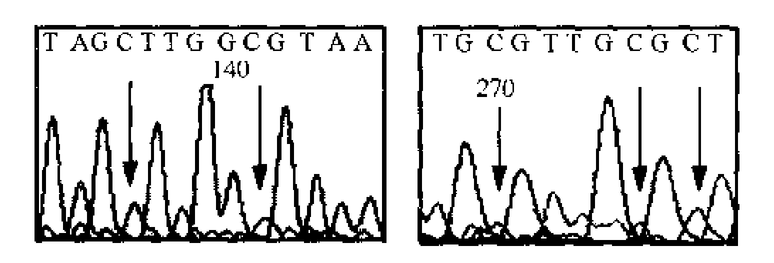


图 15-5　G 后面的 C 峰较弱

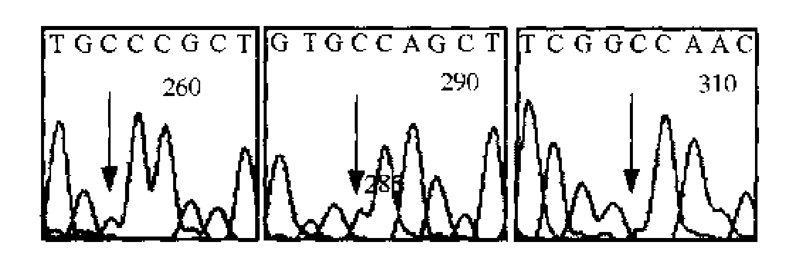


图 15-6　连续 G 峰后面的一串 C 中的第一个 C 峰较弱

(2) 碱基 G 后面 T 峰一般信号较弱（图 15-7）。

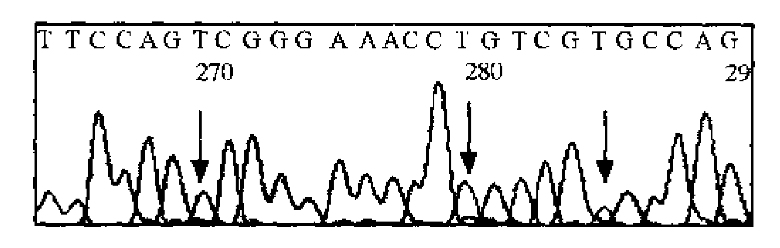


图 15-7　G 峰后面的 T 较弱

(3) 在 2 个或多个连续的 T 后面的碱基 C 峰的信号较强（图 15-8）。

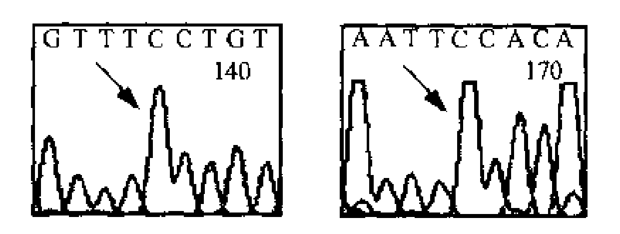
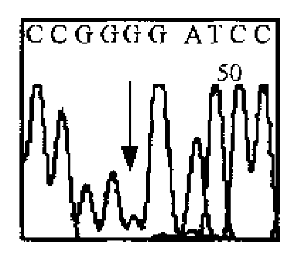


图 15-8　连续 T 峰后面 C 峰被放大

(4) 在 4 个或多个连续的 G 中，第三个 G 的信号较弱，后面的则较强（图 15-9）。



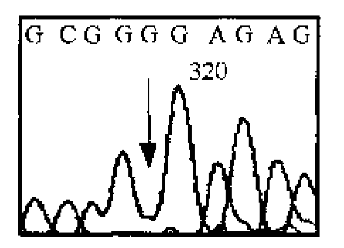


图 15-9 一串碱基 G 中的第三个碱基 G 较弱

（5）在多个连续的 T 后面的 G 峰信号较强，有时强到掩盖了附近的峰（图 15-10）。

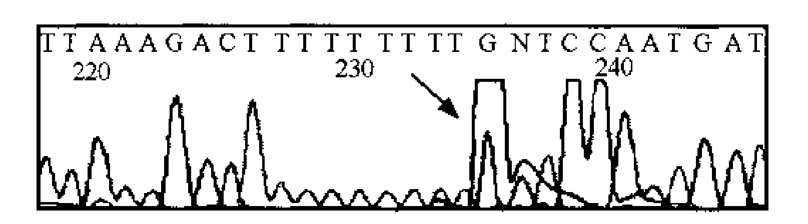


图 15-10 一串 T 后面 G 峰信号较强

（6）在多个连续的 A 中，第一个 A 峰较强，而随后的 A 峰都较弱（图 15-11）。

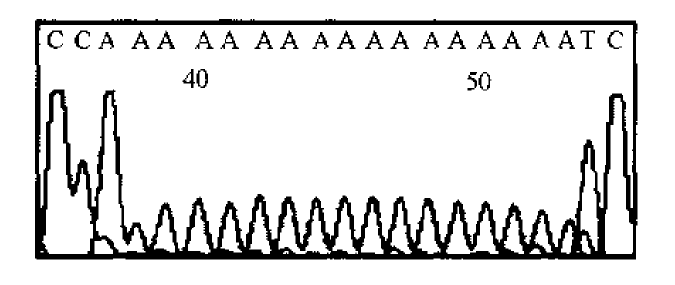


图 15-11 一串的 A 峰，第一个 A 峰较强，而随后的则较弱

（7）在多个连续的 T 中，第一个 T 峰较强，而随后的 T 峰都较弱（图 15-12）。

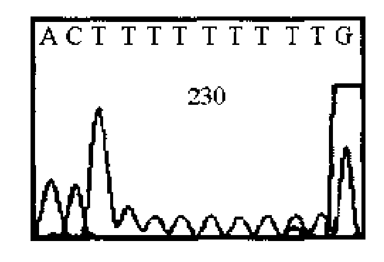


图 15-12 一串的 T 峰，第一个 T 峰较强，而随后的则较弱

（8）跟在 T 峰后而的 A 峰信号一般较弱（图 15-13）。

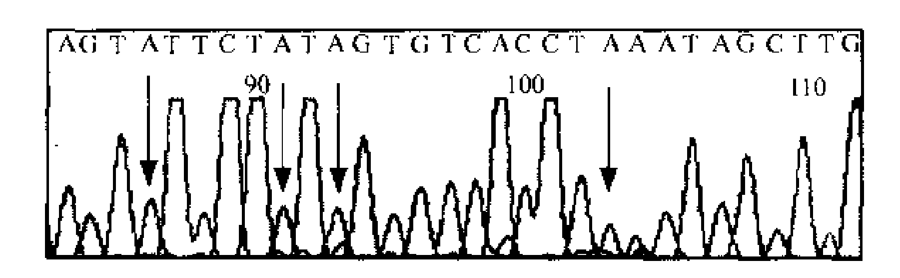


图 15-13 T 峰后面 A 峰较弱

15.4.6 荧光自动测序分析中的常见问题

1. 在 nested - PCR 中，进行二次扩增来增加测序模板的量，提高测序灵敏度，但同时也增加了污染的危险性。因此必须在每次扩增中设置阳性与阴性对照，保证测序结果的正确性。

2. 由于 PCR 反应的特异性较低或原始模板 DNA 过多，模板容易出现杂带，使测序结果杂乱无章（图 15-14）。因此，在模板准备时要求保证 PCR 反应的特异性，使 PCR 产物只有主带，而没有非特异性产物。

3. 引物浓度过低、引物反复冻融而发生降解，或由于引物浓度过高，造成序列杂带多。

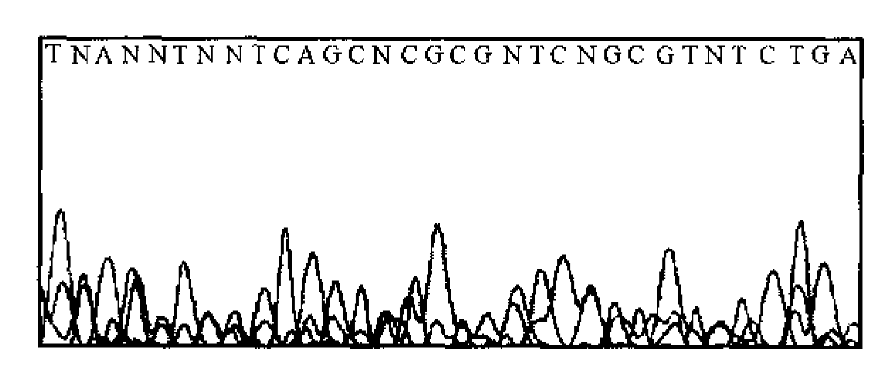


图 15-14 由于模板不纯或 PCR 引物残留导致信号杂乱

4. 在保证模板质量前提下，测序反应中的聚合酶，或核苷酸混合液交叉污染，可能引起杂带存在。

5. 与凝胶相关的问题

(1) 在凝胶制备过程中，玻璃板未洗净有油脂，或灌胶时不流畅，使凝胶中有气泡，影响片段电泳分离；

(2) 缓冲液浓度不准确，缓冲液结晶太多，电泳行为反常，片段迁移缓慢，在程序设定时间内没有分离完；

(3) 凝胶聚合不均衡，迁移不一致；

(4) 灌胶时温度太高。凝胶聚合太快，基质特性发生改变，凝胶电泳速度变慢，使峰变宽，在电泳早期分辨率下降；

（5）凝胶灌制后，放置时间过短或过长，使分辨率下降或可阅读区变短。要求凝胶聚合 2h 以上，而且在聚合后 24h 内电泳；

（6）在制备的胶液中，或玻璃板、梳子或缓冲液不小心被有荧光物质污染，在胶图中出现一片蓝色或绿色，例如，酒精、肥皂或洗涤剂，记号笔的墨水，或质量较差的甲酰胺常常是荧光污染的原因。

6. 原始数据扫描线低，无法读序，可能原因是：

（1）模板太少或质量不好；

（2）引物退火失败。测序引物浓度过低、退火温度过高，或没有引物结合点。重新设计引物序列，调整引物浓度；

（3）试剂过期或处理不当；

（4）热循环参数不当：热循环仪发生故障或温度校正有误；或循环参数不正确，或反应管与扩增仪不配套，接触不紧密；

（5）模板碱基组成影响：模板中含有大量 G－C，易形成稳定的二级结构。

7. 噪音高，序列杂乱，可能原因：

（1）模板不纯，含有抑制酶反应的杂质；

（2）多个 DNA 序列的模板，PCR 扩增特异性低，含有非特异性产物。可以用琼脂糖凝胶电泳检测是否只有一条带；

（3）在连续一串碱基处发生酶的滑动，导致跳过一个碱基或添加一个碱基；

（4）引物发生污染，或引物在多个区域发生退火，产生叠加的序列。

8. 线粒体 DNA 的 D－环高变区 1（HVRⅠ）的碱基序列中，从 16193 位置（以轻链为例）为 5 个 C，1 个 T，然后 4 个 C 即 CCCCCTCCCC，由于 polyC 的影响，随机终止，无法得到这一区域之后的数据；另一方面，由于 16189 位置存在 C/T 突变，如果随机终止，这一区域测得的碱基序列结果难以确认，通常需要进行反向测序，检测相互匹配的重链碱基序列，以补充完整结果。

9. 由于没有彻底除去未掺入到反应中的带荧光的 ddNTP，在测定的序列前 100 个碱基区域（前 40、70 及 90nt 位置）出现较宽的峰（图 15－15），在序列中会出现较高的噪音。有时会严重到这种残留碱基的峰掩盖整个样品的序列峰。有时一个较大 T 峰出现在第 220 碱基周围，无法正确阅读结果。因此在用乙醇或异丙醇沉淀测序反应产物时，必须彻底去除所有溶液，以便彻底除去未掺入的染料。

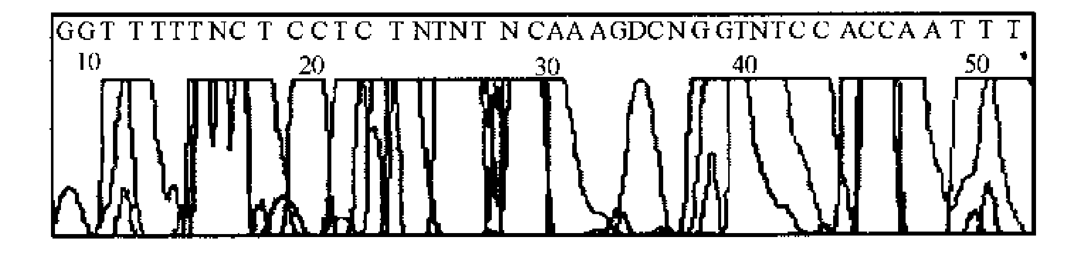


图 15－15　未掺入染料标记的 ddNTPs 峰

10. 在循环测序结果中，有时出现二或三个红色的较宽 T 峰（图 15－16）。其原因目前还不清楚。大多数情况下，小心地除去过量的染料可以消除这个问题。

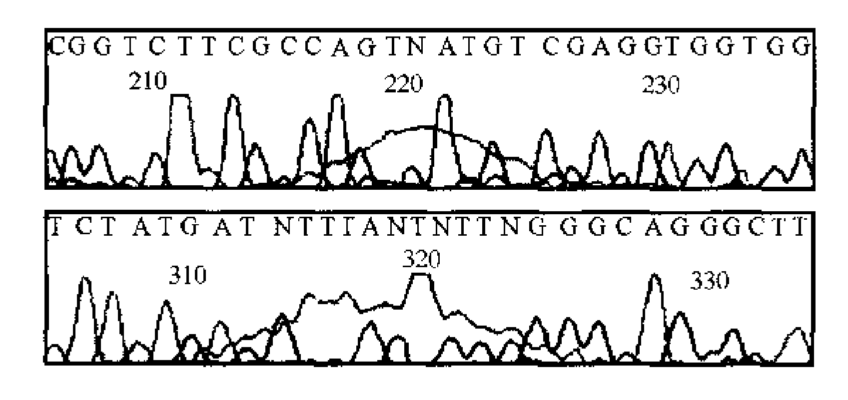


图 15－16 2 个或 3 个较宽的 T 峰

11. 在 PCR 反应过程中形成的引物二聚体在纯化时可能没有完全除去，残留的引物二聚体会影响循环测序的数据。测序引物与引物二聚体进行退火，代替目的序列，导致在序列前端 25～30 个碱基序列无法判读，合理设计引物可以避免这一问题。

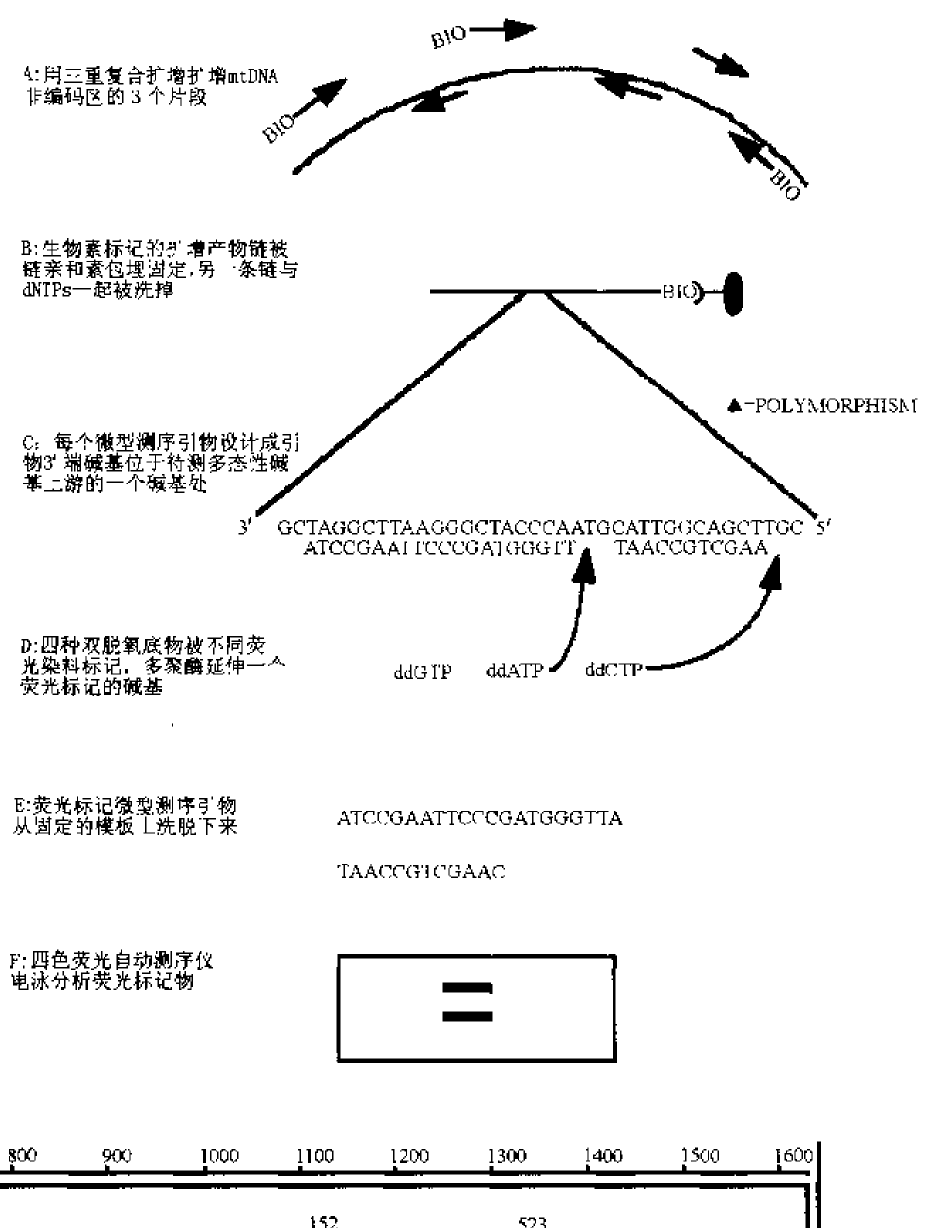
12. 信噪比：为降低噪音，获得满意的信号值，要求 PCR 反应时模板 DNA 在检测区域必须完整、高纯度和有一定的数量；PCR 产物纯化时要求回收率高，去除杂质能力强；纯化测序反应产物，尽可能彻底去除染料分子。同时，玻璃板的清洁也是十分重要。

13. 采用 PE480 扩增仪进行 PCR 反应，大约测序结果的前 40bp 不太易读。如果采用 PE2400 或 PE9600 进行 PCR 反应，可以解决这一问题。这可能是 PE480 的升降温速度较慢的原因。

14. 测序昂贵、极易发生污染等因素限制 mtDNA 在法医学应用。如 DNA 芯片技术开发更成熟，可常规应用。

15.5 四色荧光标记固相微型测序分析

1996 年 Gillian 等人报道了用复合固相微型测序检测 mtDNA D－loop 区 12 个多态性位置（H00073，H00146，H00152，H00247，H16069，H16129，H16189，H16224 和 H16311，二个长度多态性区 H00302/309 与 L00525/523），调查了英国高家索人与英国的非洲加勒比人群，发现微型测序法得到的两个无关个体的似然比分别为 0.054 和 0.026。图 15－17 为固相测序原理与步骤示意图。



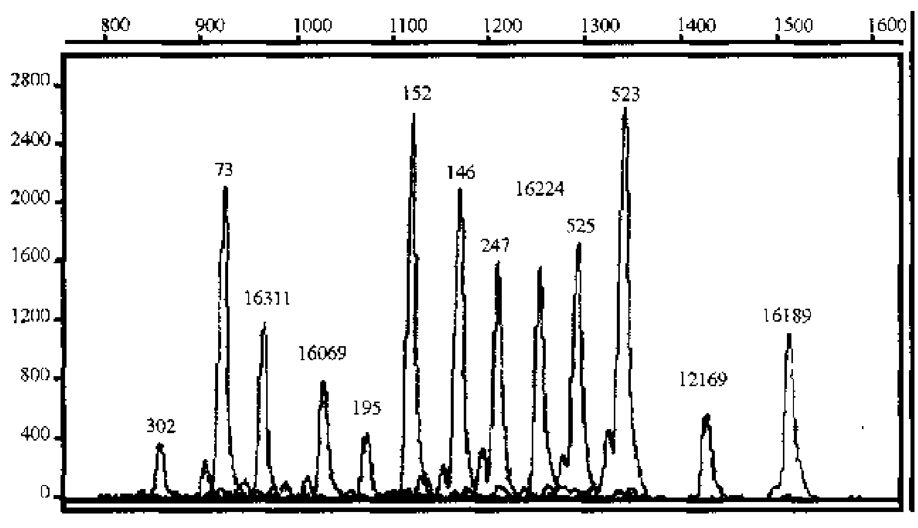


图 15-17　固相微型测序原理图（荧光标记法）

1. 生物素模板 DNA 扩增

(1) 引物　用三对引物扩增出包含线粒体 DNA 12 个多态性位置的片段，引物序列如下（5′→3′）：

L16048 生物素 - TCA TGC GAA GCA GAT TTG G

H16332 GGA TTT GAC TGT AAT GTG CTA TG

L00066 生物素 - TGC ATT TGG TAT TTT CGT CTG

H00326 CAG AGA TGT GTT TAA GTG CTG T

L00386 GAA CCC TAA CAC CAG CCT AAC

H00537 生物素 - GGA GGT AAG CTA CAT AAA CTG TG

(2) 反应体系　在 50μl 反应体系中，含有 1～10μl mtDNA，1×PCR 反应缓冲液，200μmol/L dNTPs，2.5u Taq DNA 聚合酶，0.5μmol/L 上述每种引物。扩增循环参数为 94℃ 变性 30s，57℃退火 30s，72℃延伸 90s，循环 35 次。

2. 生物素的扩增片段结合到固相支持物上

根据厂家要求将 Dynbead M - 280 用链亲和素预洗一下，重新悬浮于 2×结合和洗涤缓冲液（2mol/L NaCl，1mmol/L EDTA，10mmol/L Tris - HCl，pH7.5，简称 BW 缓冲液）中。将 40μl PCR 产物与 40μl 悬浮好的 Dynbead 在 48℃保温 15min，将 PCR 产物结合到 Dynbead 上。Dynbead - PCR 产物复合物用 100μl 2×BW 缓冲液洗，然后再用无菌蒸馏水洗结合物，最后在 0.15mol/L NaOH 溶液中室温下保温 4min 使变成单链，用 100μl 0.15mol/L NaOH 洗一次。再用 100μl 250mmol/L Tris - HCl（pH8.0）、0.1% Tween - 20 溶液洗一次，然后用 100μl 10mmol/L Tris - HCl、1mmol/L EDTA 溶液洗一次，100μl 无菌蒸馏水洗一次。洗好的珠 - 结合物悬浮在 10μl 无菌蒸馏水中。

3. 微型测序反应

(1) 微型测序反应引物　微型测序引物序列要求在多态性碱基的上游一个位置，所有测序反应的链延伸 - 终止只增加一个碱基。为了能使所有微型测序反应产物能通过凝胶电泳区分开来，引物设计时根据电泳分离需要，在引物 5′端连接多个 T，使每个引物长度不同。表 15 - 4 为测序反应引物。

表 15 - 4　测序反应引物

3′碱基定位	序列 5′ - 3′	浓度（μmol/L）
H00303	GTC GTG TGG CCA GAA GCG GGG GGA GGG GGG G	6
H0074	TTT TTT TTT CCA GCC TCT CGC AAT GCT ATC GCG TCC A	0.3
H16312	TTT TTT TTT TTT TGT GCT ATG TAC GGT AAA TGG CTT TAT GT	0.06
H16070	TTT TTT TTT TTT TTT TAA GCA TTA ATT AAT TAA CAC RCT TYR RTA	0.4
H00196	TTT TTT TTT TTT TTT TTA AGC ATT AAT TAA TTA ACA CRC TTY RRT A	6.5
H00153	TTT TTT TTT TTT TTT TTT TCT GTA ATA TTG AAC GTA GGT GCG ATA AAT AAT	0.3
H00147	TTT TTT TTT TTT TTT TTT TTA ATA TTG AAC GTA GGT GCG ATA AAT AAT RRR ATG	1.25
H00248	TTT TTT TTT TTT TTT TTT TTT TTT TTC TGT GTG GAA AGT GGC TGT GCA GAC ATT	0.05

续表

H16224	TTT TTT TTT TTT TTT TTT TTT TTT TTT TTG GAG TTG CAG TTG ATG TCT GAT AGT TG	0.1
L00522	TTT TTT TTT TTT TTT TTT TCT CAT CAA TAC AAC CCC CGC CCA TCC TAC CCA GCA CAC ACA CAC	0.6
L00524	TTT TTT TTT TTT TTT TTT TTT TTT CTC ATC AATACA ACC CCC GCC CAT CCT ACC CAC CAC ACA CAC	0.48
H16130	TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TGT ACT ACA GGT GGT CAA GT	0.04
H16190	TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT GGT TGA TTG CTG TAC TTG CTT GTA AGC ATG RGG	0.2

（2）微型测序反应　在 0.2ml 薄壁离心管中，加入 40μl 含有 0.4mol/L Tris－HCl（pH9.0），10mmol/L $MgCl_2$，20%二甲基硫氧化物（dimethylsulphoxide），0.1μl 荧光标记的 ddCTP 和 ddTTP，0.125μl 荧光标记的 ddATP 和 ddGTP，2.5u Taq FS 和 3μmol/L 上述 13 个引物，PE 公司 9 600 扩增仪预先预热到 57℃，将含有测序反应液的薄壁离心管和含有珠－产物复合物的管子放入扩增仪中，57℃30s 后，将测序反应液加入珠－产物复合物的管子中，52℃下保温 60s，然后将珠－产物结合物迅速放入冰水浴中，将测序混合液吸去。

4. 电泳分离

检测样品重新悬浮于 100μl 1×BW 缓冲液中。去掉 1×BW 缓冲液后，样品悬浮于 4μl 甲酰胺中，72℃保温 4min，然后迅速放入冰水浴中。取 1μl 加样在 377 测序仪的 19%变性聚丙烯酰胺凝胶中，1 680V 电泳 2h，分析数据。

5. 结果分析

根据凝胶电泳图片中谱带颜色与出现的位置确定核苷酸，得到单倍型。图 15－17 为线粒体 DNA 的 12 个多态性位置的结果，图谱相对简单，可以进行混合样品分析。

使用底物四色荧光标记测序，可以同时检测出取代的每一种碱基类型，比同位素或非同位素标记省事、省时，费用低，更加准确可靠。多个多态性点可以一起检测分析，大大提高了效率。

15.6 法医学应用

15.6.1 意义

1. 由于线粒体 DNA 的高突变率，线粒体 D－loop 区存在着三个高变区（HVR Ⅰ、Ⅱ和Ⅲ），具有高度个体差异性，尤其是高变区Ⅰ，同一人群不同个体间存在差异，人群与人群也存在差异，比如在非洲裔美国人群个体间平均有 14 个核苷酸不同，高家索人群中只有 8 个核苷酸不同。通过对 105 名中国汉族无关个体的 mtDNA D－loop 区 HVR Ⅰ（15997～16401）测序分析，发现存在 118 个碱基突变点，各位置的突变率为表 15－5，在此区域，不同个体平均突变碱基数为 4.029 个，最多可达 9 个碱基突变。序列分析两个高变区，排除率可达 95.6%～98.1%，线粒体 DNA 具有高度序列多态性，可以用于个体识别。

表 15－5 mtDNA15997～16401 之间突变位置的碱基分布频率

碱基号	A（%）	C（%）	G（%）	T（%）
16016	99.048			0.952
16019		99.048	0.952	
16020		0.952		99.048
16021		97.143	2.857	
16022			0.952	99.048
16023			99.048	0.952
16025		0.952		99.048
16027		0.952		99.048
16028	0.952			99.048
16030		93.333	6.667	
16037	98.095		1.905	
16042	0.952		99.048	
16050	0.952			99.048
16053		99.048		0.952
16055	0.952	99.048		
16060	0.952		99.048	
16061		0.952		99.048
16068		1.905		98.095
16069		99.048	0.952	
16075		0.952		99.048
16085		96.190		3.810
16086		4.772		95.228
16092		3.810		96.190
16093		7.619		92.381
16094			0.952	99.048
16095		99.048	0.952	
16102		0.952		99.048
16107		98.095		
16111		95.238		4.762
16124		0.952		99.048
16126		0.952		99.048
16129	26.667		73.333	
16134		98.095		1.905
16136		3.810		96.190
16139	99.048		0.952	
16140		2.857		97.143
16144		1.905		98.095
16145	3.810		96.190	
16149		99.048		0.952
16152	1.905			98.095
16158	99.048		0.952	

续表

碱基号	A（%）	C（%）	G（%）	T（%）
16161		0.952		99.048
16162	96.190		3.810	
16169		96.190		3.810
16172		93.333		6.667
16179		97143		2.857
16182	98.095	1.905		
16183	97.143	2.857		
16184		94.286		5.714
16185		92.382		7.619
16187		98.095		1.905
16188		99.048		0.952
16189		7.619		92.382
16192		98.095		1.905
16197		0.952		99.048
16209		3.810		96.190
16217		2.857		97.143
16223		28.571		71.429
16227	97.143		2.857	
16230	98.095		1.905	
16233	97.143		2.857	
16234		93.33		6.67
16235	98.095	0.952	0.952	
16239		99.048		0.952
16240	99.048	0.952		
16249		0.952		99.048
16255	0.952		99.048	
16256		95.238		4.762
16257	4.762	95.238		
16259		99.048		0.952
16260		95.238		4.762
16261		93.333		6.667
16262	2.857	97.143		
16264	0.952	99.048		
16266		94.286		5.714
16268		99.048		0.952
16274	3.810		96.190	
16278		96.190		3.810
16288		0.952		99.048
16290		92.381		7.619
16291		96.190		3.810
16292		98.095		1.905

续表

碱基号	A（%）	C（%）	G（%）	T（%）
16293		99.048		0.952
16295		99.048		0.952
16296		99.048		0.952
16297		7.619		92.381
16298		8.571		91.429
16299	98.095		1.905	
16300	99.048		0.952	
16304		13.333		86.667
16308		0.952		99.048
16311		5.714		94.286
16312	99.048		0.952	
16316	99.048		0.952	
16317	98.095		1.905	
16319	7.619		92.381	
16324		0.952		99.048
16325		3.810		96.190
16327		95.238		4.762
16334		0.952	99.048	
16335	97.143		2.857	
16337		99.048	0.952	
16347			0.952	99.048
16354		99.048		0.952
16358		3.0810		96.190
16359		99.048		0.952
16360		3.810	96.190	
16361		0.952		99.048
16362		9.524		90.476
16367	99.048	0.952		
16371	99.048			0.952
16376	0.952		99.048	
16374	97.143		2.857	
16376		99.048	0.952	
16385	99.048			0.952
16390	0.952		99.048	
16393		94.280	5.714	
16395		99.048		0.952

2. 线粒体呈母系遗传，兄弟姐妹、婊兄弟姐妹、姨的 mtDNA 序列一致，已有实验数据表明四代之内，所有母系亲属的线粒体序列相同，可以进行母子（女）鉴定、身源鉴定，如 Holland（1993）用 mtDNA 测序法对越战身亡的美国士兵残骸进行了身源鉴定，尤其是那些无法得到父母 DNA 样品进行鉴定的情况下，利用身亡士兵兄弟的 mtDNA 序列进

因此，检验时不能单凭一个碱基序列的差异轻易排除其相似性，3 个以上碱基不同排除较为可靠。同一母系人员间序列也存在异质性，在亲缘鉴定时要予以注意。英国的 FSS 和美国的 FBI 及军队武装 DNA 鉴定中心规定不能根据单个碱基不同而得出排除结论。

5. mtDNA 序列单倍型数多，但是现在调查的人群数据少，还未获得真正有代表性的单倍型的群体数据，很难进行个体同一认定和亲子鉴定的概率计算。

6. mtDNA 序列分析方法复杂、繁琐。工作量大，费用昂贵，出错率高。Holland 和 Parsons 报道，mtDNA 测序错误率为 1% ~ 2%。

15.6.3 分型结果解释

进行物证样品比对时，首先要检查提取试剂对照或 PCR 阴性对照序列。如果提取试剂对照或 PCR 阴性对照得到与物证样品相同的序列，那么物证样品的结果绝对不能使用，必须重新分析。如果对照样品出现的序列与物证样品的不同，结果解释时必须结合所有实验结果，包括阳性对照、其他阴性对照、两次分析的结果和阴性对照相对于物证样品的相对强度。每次实验的阳性对照的结果必须是正确的，本次实验结果方可应用。

在确保 mtDNA 分型正确情况下，比较物证样品与参照样品序列。根据人群调查发现，无关个体在高变区Ⅰ和Ⅱ一般有 4 个以上碱基序列不同，异质性很少同时发生在两个碱基中，不是来自于同一母系的那些物证与比对样品间不止在一个核苷酸上有差异。

1. 如果比对样品有 2 个以上的碱基序列不同，可以排除他们的相关性。

2. 如果比对样品序列只有一个碱基差异。分两种情况：(1) 如果物证样品观察到异质体，而没有在已知样品中观察到，或反过来，在已知样品中观察到，而在物证样品中没有观察到，这种情况下不能排除他们有同一母系关系或来自于同一来源。(2) 如果未知样品和比对样品间只有一个碱基差异，而又没有观察到异质性的存在，此时的结果解释应是不能确定（inconclusive）。已发现在个别同一母系人员间存在一些差异，有时不同母系或来源个体出现单核苷酸不同。

总之，样品间碱基差异数越大，他们来自于同一来源或有相同母系起源的可能性越小。

3. 如果序列相同，那么不能排除物证样品与参照样品来自于同一个来源的可能。比如现场提取的毛发与某一嫌疑人 A 的相同，那么这种匹配有三种可能情况：(1) 毛发可能来自于嫌疑人；(2) 毛发可能来自与嫌疑人有同一母系的亲属；(3) 毛发可能来自与嫌疑人无母系关系的无关个体。

如何判定结果属于哪种情况？属于其中第三种情况的可能性有多大？也就是说，具有这种毛发序列的潜在人群占多少？这一般取决于所分析的案件的案情以及所涉及到的 mtDNA 序列。如果案件限定在一定范围内，那么 mtDNA 序列相匹配可以认为是有决定意义的。如两个比对样品中同时观察到相同的异质体，这就增加了物证的证据力。又如，在飞机失事的空难中，采集的残骸是登机的乘客与机组成员的，也就是说，mtDNA 序列比对局限在一个比较固定有限的范围内。如果某一乘客的参考样品（乘客的母系家属）与他的序列相同，但与所有其他人员的序列不同，那么这种相匹配可以认为是同一的。

在实际工作中，更常见的是，进行 mtDNA 分析的案件物证的来源是未知的，没有任何关联，那么在这种情况下的 mtDNA 序列相同的资料是不能作为最终结论的决定性依据。

因此，检验时不能单凭一个碱基序列的差异轻易排除其相似性，3 个以上碱基不同排除较为可靠。同一母系人员间序列也存在异质性，在亲缘鉴定时要予以注意。英国的 FSS 和美国的 FBI 及军队武装 DNA 鉴定中心规定不能根据单个碱基不同而得出排除结论。

5.mtDNA 序列单倍型数多，但是现在调查的人群数据少，还未获得真正有代表性的单倍型的群体数据，很难进行个体同一认定和亲子鉴定的概率计算。

6.mtDNA 序列分析方法复杂、繁琐。工作量大，费用昂贵，出错率高。Holland 和 Parsons 报道，mtDNA 测序错误率为 1% ~ 2%。

15.6.3 分型结果解释

进行物证样品比对时，首先要检查提取试剂对照或 PCR 阴性对照序列。如果提取试剂对照或 PCR 阴性对照得到与物证样品相同的序列，那么物证样品的结果绝对不能使用，必须重新分析。如果对照样品出现的序列与物证样品的不同，结果解释时必须结合所有实验结果，包括阳性对照、其他阴性对照、两次分析的结果和阴性对照相对于物证样品的相对强度。每次实验的阳性对照的结果必须是正确的，本次实验结果方可应用。

在确保 mtDNA 分型正确情况下，比较物证样品与参照样品序列。根据人群调查发现，无关个体在高变区Ⅰ和Ⅱ一般有 4 个以上碱基序列不同，异质性很少同时发生在两个碱基中，不是来自于同一母系的那些物证与比对样品间不止在一个核苷酸上有差异。

1. 如果比对样品有 2 个以上的碱基序列不同，可以排除他们的相关性。

2. 如果比对样品序列只有一个碱基差异。分两种情况：（1）如果物证样品观察到异质体，而没有在已知样品中观察到，或反过来，在已知样品中观察到，而在物证样品中没有观察到，这种情况下不能排除他们有同一母系关系或来自于同一来源。（2）如果未知样品和比对样品间只有一个碱基差异，而又没有观察到异质性的存在，此时的结果解释应是不能确定（inconclusive）。已发现在个别同一母系人员间存在一些差异，有时不同母系或来源个体出现单核苷酸不同。

总之，样品间碱基差异数越大，他们来自于同一来源或有相同母系起源的可能性越小。

3. 如果序列相同，那么不能排除物证样品与参照样品来自于同一个来源的可能。比如现场提取的毛发与某一嫌疑人 A 的相同，那么这种匹配有三种可能情况：（1）毛发可能来自于嫌疑人；（2）毛发可能来自与嫌疑人有同一母系的亲属；（3）毛发可能来自与嫌疑人无母系关系的无关个体。

如何判定结果属于哪种情况？属于其中第三种情况的可能性有多大？也就是说，具有这种毛发序列的潜在人群占多少？这一般取决于所分析的案件的案情以及所涉及到的 mtDNA 序列。如果案件限定在一定范围内，那么 mtDNA 序列相匹配可以认为是有决定意义的。如两个比对样品中同时观察到相同的异质体，这就增加了物证的证据力。又如，在飞机失事的空难中，采集的残骸是登机的乘客与机组成员的，也就是说，mtDNA 序列比对局限在一个比较固定有限的范围内。如果某一乘客的参考样品（乘客的母系家属）与他的序列相同，但与所有其他人员的序列不同，那么这种相匹配可以认为是同一的。

在实际工作中，更常见的是，进行 mtDNA 分析的案件物证的来源是未知的，没有任何关联，那么在这种情况下的 mtDNA 序列相同的资料是不能作为最终结论的决定性依据。

此时必须考虑到这种 mtDNA 序列在人群中的分布频率。根据数据库内的有关 mtDNA 单倍型分布数据和相关群体的亚群体资料，特定序列或单倍型出现的频率、采用一些样品错误的校正值（如置信区间，bootsrtap 等），给出估计的 mtDNA 单倍体型的频率。例如，英国内务部法庭科学服务中心收集了 163 个高家索人线粒体 DNA 序列，其中有 11 人拥有同一个序列，即 7% 人群拥有相同一种序列，这种匹配中，其匹配的意义就很低。如果比对的序列是第一次发现，在数据库中没有相同序列记录在案，也就是说，这种序列在一般人群中较少观察到，那么此时的匹配就表明现场样品很可能来自嫌疑人或嫌疑人的同一母系亲属。

第十六章　SNP（单核苷酸多态性）分析

16.1 概述

SNP（single nucleotide polymorphism）是单核苷酸多态性的简称，指基因组内特定核苷酸位置上存在有两种不同的碱基。SNP 其实是一种碱基序列多态性，对它的研究开始于 20 世纪 80 年代，只是那时候没有把它叫成 SNP，而是通俗地称为序列多态性，或叫点突变。比如 HLA-DQα，就是一个典型的 SNP 标记。后来开发的 Polymarker 系统，一些血型和蛋白酶型的 DNA 分型系统都是 SNP。直到 20 世纪 90 年代末期，科学家们提出 SNP 这个专业名词，来描述碱基序列的多态性，进行更广泛的研究、开发与应用。

16.1.1 SNP 形成原因

碱基替换以及单个碱基的插入或缺失导致某一位置碱基发生变化，表现出一种二等位基因标记（biallelic markers）（或二态遗传变异）、三等位基因标记或四等位基因标记，其中以二等位基因标记较常见，形成一个 SNP。在 SNPs 基因座，碱基替换大多数为转换，即一种嘌呤替换成另一种嘌呤或一种嘧啶替换成另一种嘧啶，少数为颠换，即嘌呤与嘧啶之间的互换。替换与颠换的比例为 2∶1。研究发现，SNPs 在 GC 序列上较多见，而且多是由 C→G。

16.1.2 SNP 的特点

（1）SNP 广泛存在于基因组中，大约每 1 000bp 存在一个 SNP，在人基因组中大约有 300 万个 SNP，其数量要比 STR 基因座超出几个数量级。SNP 不仅广泛存在于非编码区，而且存在于编码区，SNP 在个体识别与亲子鉴定中具有巨大的潜力。

（2）SNP 大都表现为二等位基因标记，即人群中只有二个等位基因和三种基因型。易于分型和确定基因频率。很适合混合样品的分析。

（3）SNP 基因座的片段更短，更适合 PCR 扩增，更能分析降解 DNA，因为引物可以设计得很近；而且不会出现像 STR 基因座因聚合酶的滑脱产生的阴影带问题。

（4）SNP 分析方法较多，有些不涉及电泳技术和片段长度测定等一系列技术问题，其中一些技术更适合于自动化。

一些学者认为，SNP 是第三代遗传标记。表 16-1 是 SNP 与 STR 的比较。

16.1.3 法医学应用

SNP 标记可以应用于个体识别和亲权鉴定。3～4 个相邻的这种标记构成的单倍型可达 8～16 种，相当于一个微卫星标记形成的多态性。理论上 50 个这样的 SNP 标记可以区分开目前全球所有无关个体，累积识别率可以达到与 DNA 多基因座指纹图一样高的水平。根据它的无比优越性和潜力也必将应用在法医物证鉴定领域里，在解决现有的 DNA 分析方法的不足，特别是提高对降解 DNA 的分析能力，以及根据现场生物物证的 DNA 基因分析，来构建 DNA 来源人的特征等对侦查破案非常有用的问题上，将会有重要的意义，可

以说这将是法医 DNA 分析的发展趋势。单核苷酸多态性虽然发现和应用才短短几年，但已经应用到研究各种生物包括人的基因表达、个体识别等许多方面。

表 16-1　SNP 与 STR 的比较

特征	STR	SNP
人类基因组中存在情况	约每 15kb 出现 1 个	约每 1kb 出现 1 个
信息量	高	低，大约为 STR 的 20%～30%
标记类型	二、三、四聚串联重复标记	二等位基因标记
每个标记的等位基因数	5 个以上	一般 2 个
检测方法	凝胶或毛细管电泳	序列分析方法，微芯片杂交
复合能力	用多色荧光技术可以把 10 多个标记可以复合在一起	微芯片可以同时检测 1 000 多个标记

16.1.4 法医学应用需要的 SNP 数目

每个 SNP 基因座只有二个等位基因，其识别率是比较低的，法医学进行个体识别、亲子鉴定需要多少个 SNP 才能达到目前 STR 基因座的检验水平？Gill（2001）认为 50 个 SNP 完全可以达到 12 个 STR 基因座的识别率。Ranajit（1999）等人提出 25～45 个 SNP 点可以达到 CODIS 系统的 13 个 STR 基因座的随机匹配概率（图 16-1），当每个 SNP 的等位基因频率为 0.1 与 0.9 时，62 个 SNP 的平均匹配概率为一百亿分之一。对于有母子参加的亲子鉴定，SNP 的等位基因频率为 0.5 和 0.1 时，需 33～81 个 SNP 才能达到 99.9%的排除概率（16-2a），而对于只有孩子参加而没有母亲参与的亲子鉴定，这些 SNP 只能达到 80%的排除率（图 16-2b）。

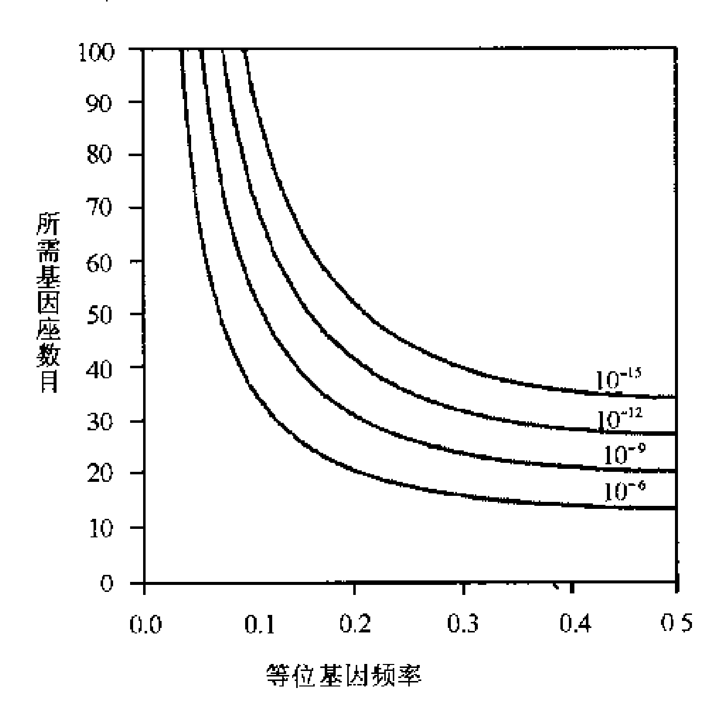


图 16-1　相当于 13 个 STR 基因座识别率的 SNP 基因座数

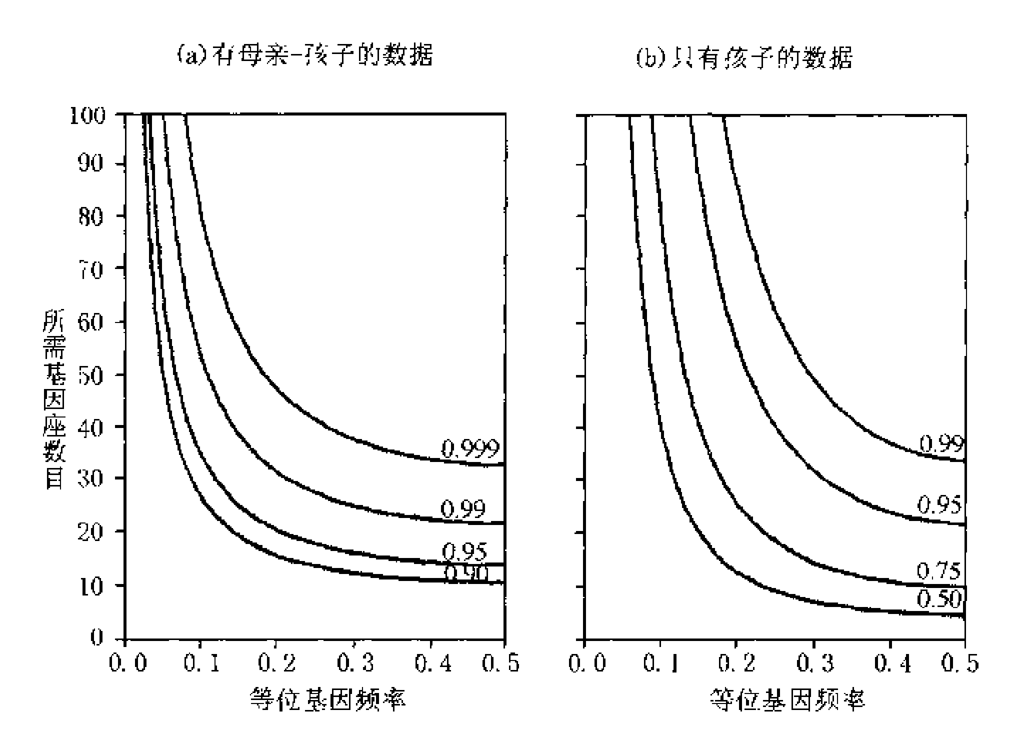


图 16-2 亲子鉴定所需的 SNP 位点数

(a) 母子参加；(b) 只有孩子参与鉴定

16.2 SNP 的分析方法

SNP 的分析检验方法很多，但也主要以 PCR 为基础，利用碱基差异造成的 DNA 片断的各种差异来进行基因分型。如利用单碱基差异引起限制性内切酶切点的产生或缺失进行限制性内切酶酶切分析，或者利用这个差异碱基设计等位基因特异引物进行特异扩增，或利用单个碱基的差异导致的构象不同进行构象分析。早期常用的方法有：利用限制性内切酶切点进行 PCR 扩增产物的 RFLP（PCR-RFLP）分析，正向或反向特异性寡核苷酸探针杂交（ASO-PCR）分析，单链构象多态性（SSCP）分析，等位基因特异引物扩增分析，异源双链分析，变性梯度凝胶电泳（DGGE）分析，直接测序和连接酶反应（ligase oligonucleotide analysis，LOA）等。DNA 直接测序分析方法复杂、成本高；SSCP 等技术难度高、信息量少、操作复杂，以上所有方法的最大缺陷是费时、信息量少。目前已经建立了多种 SNP 分析新方法，如：根据因等位基因序列不同产生的熔解温度差异进行 Tm 漂移（Tm-shift）分析，变性高压液相色谱（DHPLC）分析，分子信号（molecule beacons）分析，引物延伸结合时间飞行质谱（time-of-flight mass spectrometry，TOF-MS）分析，微排列和 DNA 芯片（DNA chip）分析等。表 16-2 列出了部分 SNP 的分析方法。本节仅就一部分方法进行介绍。

表 16－2 部分 SNP 分析方法

方法	简介
Genetic bit analysis	用 ddNTP 进行引物延伸，颜色反应检测
直接测序	PCR 产物进行测序，比较分析发现 SNP 点
变性高压液相色谱（DPHLC）	混合二个 PCR 产物，进样到离子对反相 HPLC，如果二个扩增产物的单个碱基存在差异，可观察到额外的异源双链峰
TaqMan 5′－核酸酶分析	由检测和淬灭染料组成的荧光探针加到 PCR 反应中，探针特异的扩增产物引起探针的酶切，荧光信号增加
荧光极性化	ddNTP 为底物对 SNP 点进行引物延伸，检测荧光素的极性变化来显示染料结合到引物上
质谱	ddNTP 作底物对 SNP 点进行引物延伸，测量引物和延伸产物的质谱差异来确定核苷酸位置
高密度阵列	在微芯片上排列成千上万个寡核苷酸探针，荧光标记的 PCR 产物与之杂交，结合到互补的探针上，显示 SNP
电子点杂交	所有已知的 SNP 等位基因排列在芯片上，用电子杂交，控制杂交强度，确定 SNP
分子信号（molecular beacons）	发卡结构的寡核苷酸探针含有一个荧光分子和淬灭分子，当淬灭分子与 DNA 目标序列杂交结合时，发出荧光，如果两个探针中存在有互补碱基，二个探针发生连接，化学显色反应检测
熔解温度曲线基因分型	进行等位基因特异 PCR，GC－尾巴连接到等位基因特异的正向链引物，带有 GC－尾巴引物扩增的等位基因在较高温度出现熔解温度曲线。

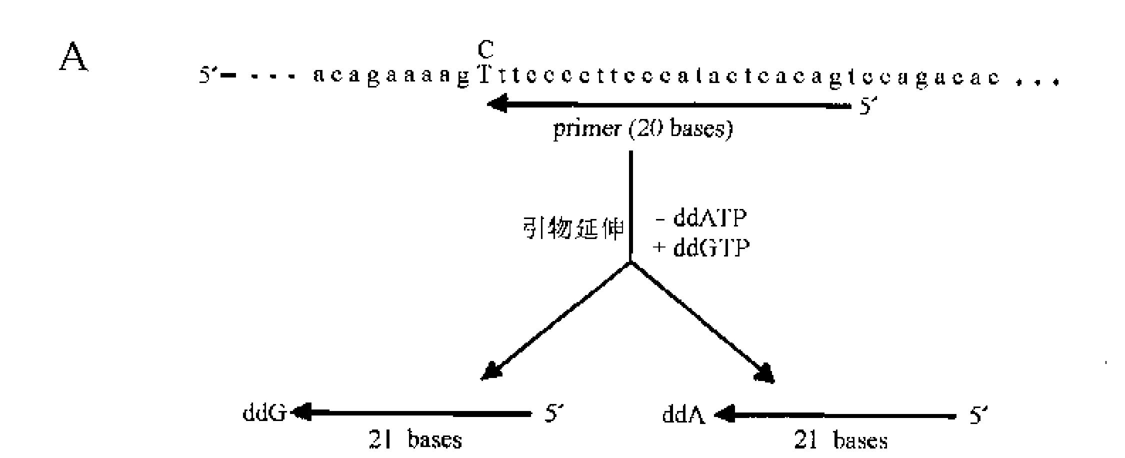
16.2.1 引物延伸结合变性高压液相色谱法

1. 引物延伸法原理

引物延伸法原理类似于微型测序方法，首先扩增出含有 SNP 的一段 DNA，然后一个寡核苷酸引物直接在多态性碱基上游或下游发生退火，在底物 dNTP 和 ddNTP 存在下，进行一个或几个碱基的延伸。延伸的碱基数目取决于多态性位置的碱基（图 16－3）。寡核苷酸引物的 3′端碱基紧挨于多态性碱基，引物开始延伸的第一个碱基就是多态性碱基。然后根据延伸产物的大小确定等位基因。关于引物延伸产物的检测，已报道了许多方法，如放射性同位素标记法、发光检测法、显色反应，凝胶为基础的荧光检测法和质谱分析法、变性高压液相色谱法。

2. 变性高压液相色谱原理

变性高压液相色谱（denaturing high performance liquid chromatography，DHPLC）又叫温度调控高效液相色谱（temperature－modulated high performance liquid chromatography，TmHPLC），是一种新的高通量的 DNA 序列多态性分析技术。该技术最早由美国 Stanford 大学 Oefner 及 Underhill 等报道的，其原理是用离子对反向高压液相色谱分离检测异源双链。优点是加样和分析的自动化，引物延伸后不需要进行纯化过程，是比较简便的 SNP 分析方法，而且多个 SNP 可以复合在一起。



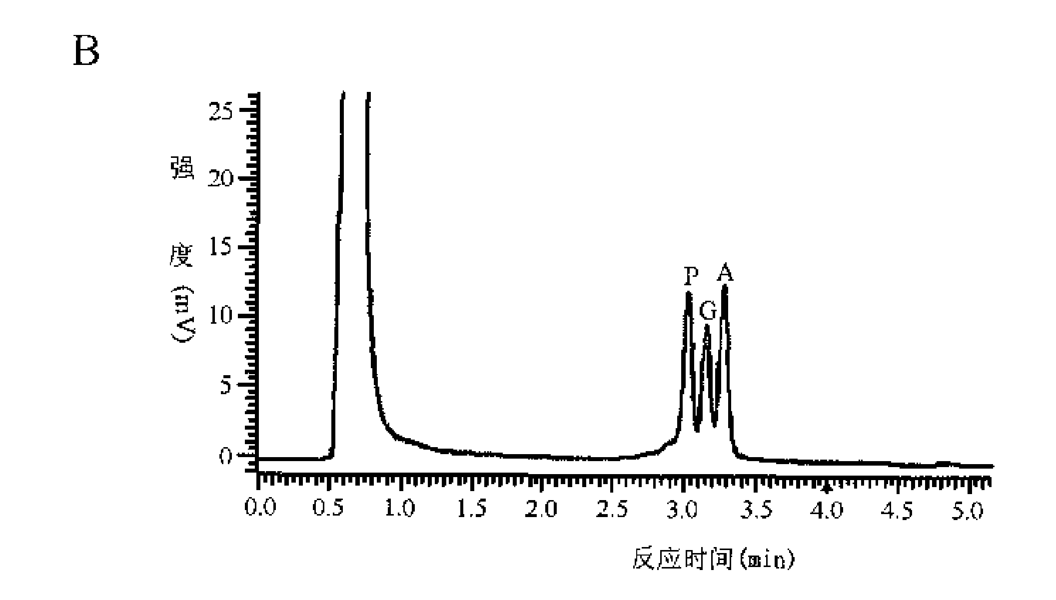


图 16－3 引物延伸法的原理示意图

根据存在的等位基因，引物可以只延伸一个碱基或延伸3个核苷酸，然后根据延伸产物大小分离等位基因。

3．DPHLC 分析 SNP

取 15μl 引物延伸产物，用 Dynamax 自动加样器（AI－96）在 70℃下加到 SarSep DNAsep 柱（Transgenomic，Santa Clara，Calif，USA）上，用含乙晴线性梯度的 0.1mol/L TEAA（pH7.0）洗脱将引物延伸产物从柱上洗脱下来，洗脱梯度由 0.1mol/L TEAA（pH7.0）＋0.1mmol/L EDTA－Na_2 与 0.1mol/L TEAA（pH7.0）＋25％乙晴混合而成。TEAA 的作用是与 DNA 分子形成离子对，使 DNA 分子呈电中性，吸附在经硅烷基化处理的色谱柱上。乙晴作为洗脱剂，将 TEAA 和 DNA 分子复合体从色谱柱上洗脱下来。

根据 DNA 片段大小和片断单链的组成顺序从柱上洗脱下来，以峰的形式被记录下来，第一个峰为掺入的引物峰，然后是引物延伸产物峰，根据峰的个数和对称性，确定基因型。杂合子个体出现二个峰，但在分离度不够的情况下，往往出现肩峰及不对称峰。

16.2.2 引物延伸结合时间飞行质谱分析法

1　原理

引物延伸结合时间飞行质谱分析（time - of - flight mass spectrometry，TOF - MS）的原理如图 16 - 4 所示。利用 PCR 技术扩增出含有多态性碱基的产物，然后用磷酸酶处理 PCR 产物以除去未参加反应的底物 dNTP。消化产物进行引物延伸。在延伸反应中，延伸引物 5′端含有一生物素分子，在 3′端有酶切位点，而且引物 3′端在 SNP 的上游一个碱基处，使延伸反应开始掺入的第一个碱基为双脱氧底物，终 [illegible]

反应。SNP 引物延伸分析中，采用 5′端含有生物素和 3′端附近有酶切位点的引物，与 SNP 位置上游退火结合。引物 5′端生物素用来在 MS 分析前固相捕捉纯化样品，3′端附近酶切点允许引物 3′端从固定部分释放出来，同时也降低 DNA 分子的质量。SNP 位置互补的核苷酸在延伸反应过程中被插入，由于在延伸过程中只有 ddNTPs 底物，延伸过程中引物只添加一个碱基。对于纯合子个体，产生一个延伸产物，对于杂合子，产生二个延伸产物。如果延伸反应不彻底，还剩余部分引物，在引物延伸结束后有引物和延伸产物同时存在。用 MS 分析这二个 DNA 寡核苷酸的质量差异来确定 SNP 位置存在的核苷酸。用这个方法可以确定 4 种可能的 SNP 纯合子 和 6 种可能的杂合子。

平衡引物延伸反应使反应中剩余一些引物，用作内部标准，提高质量差异测定的准确性和精确性，保证正确命名 SNP 核苷酸。试验观察到的每个脱氧核苷酸的质量分别为：ddC 为 273.5 ± 0.5，ddT 为 288.3 ± 0.7，ddA 为 297.2 ± 0.5，ddG 为 313.2 ± 0.5。图 16 – 5 展示了 50 个 SNP 的 4 个可能的碱基的测量准确性和精确性。

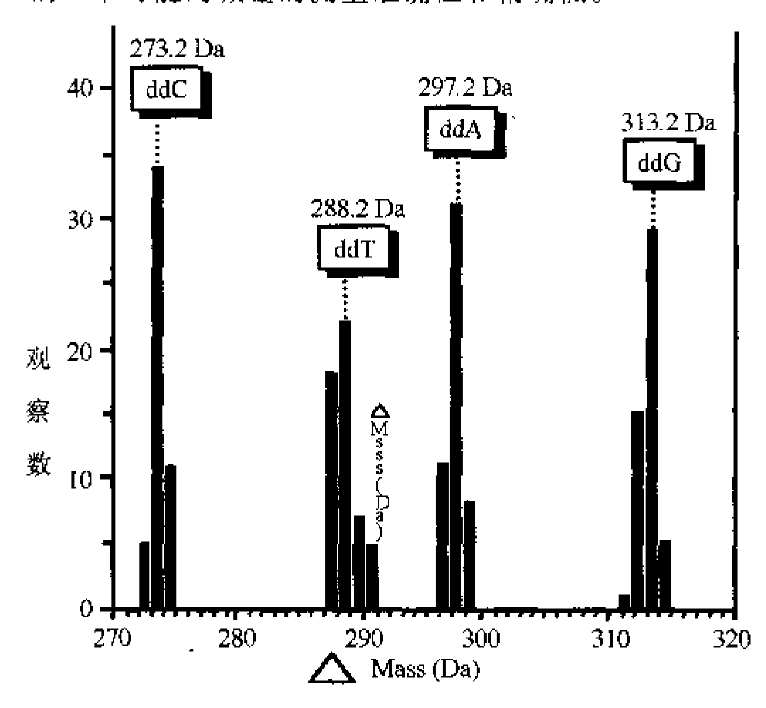


图 16 – 5 200 个样品的质量差异图

双脱氧核苷酸的期望质量分别是：ddC 273.2，ddT 288.2，ddA 297.2，ddG 313.2。

从分辨率角度看，最难以区分开的是碱基 A 与 T，它们相差只有 9 Da（如图 16 – 6），可以通过限制性内切酶酶切引物方法，使二者长度差异更大些，解决 T/A 杂合子。含有 C 和 G 的杂合子最容易分离，它们相差 40 Da。

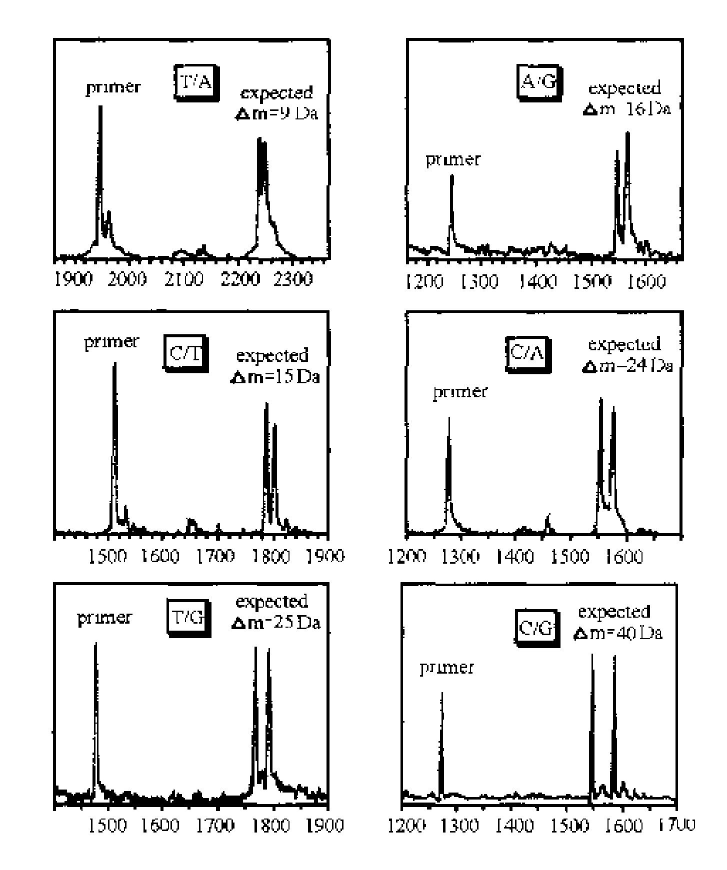


图 16－6 杂合子 SNP 的质量谱

列出了 ddNTP 的 6 个可能的组合（T/A，C/T，A/G，C/A/T/G 和 C/G）与它们期望的质量差异。测量引物和延伸产物间的质量差异确定每个样品的基因型。表 16－3 是这些杂合子样品测得的质量资料。

表 16－3 杂合子 SNP 检测结果

引物序列	期望的引物质量（Da）	引物质量（Da）	延伸质量（Da）	质量差异（Da）	SNP 位点	延伸产物之间的质量差（Da）
GACGGG	1937	1942.0	2231.0	289.0	T/A	9.1
			2240.1	298.1		
CAAA	1247	1244.7	1541.7	297.0	A/G	18.4
			1560.1	315.3		
ACACA	1536	1538.6	1812.9	274.3	C/T	15.1
			1828.0	289.4		
CGAG	1279	1277.7	1550.8	273.1	C/A	23.2
			1574.0	296.3		
TTGTC	1485	1490.5	1780.4	289.9	T/G	25.0
			1805.4	314.9		
GAAT	1278	1275.2	1547.8	272.6	C/G	39.9
			1587.7	312.5		

3. 多个 SNP 标记复合

根据不同质量的引物可以同时检测多个标记复合，在同一个反应内检测多个 SNP 标记，节省时间、人力和费用。对于退火温度近似的引物，在每个引物不同位置设置酶切点。例如一个引物在 3'端第 5 个碱基处引入酶切点，而另一个引物的酶切点与 3'端之间有 8 个碱基。因此在 SNP 延伸反应和酶切引物后，根据质量参数很容易区分开每个引物和延伸产物。图 16－7 为 6 个线粒体 DNA SNP 复合结果。表 16－4 为这些人类线粒体 DNA SNP 标记的具体资料。

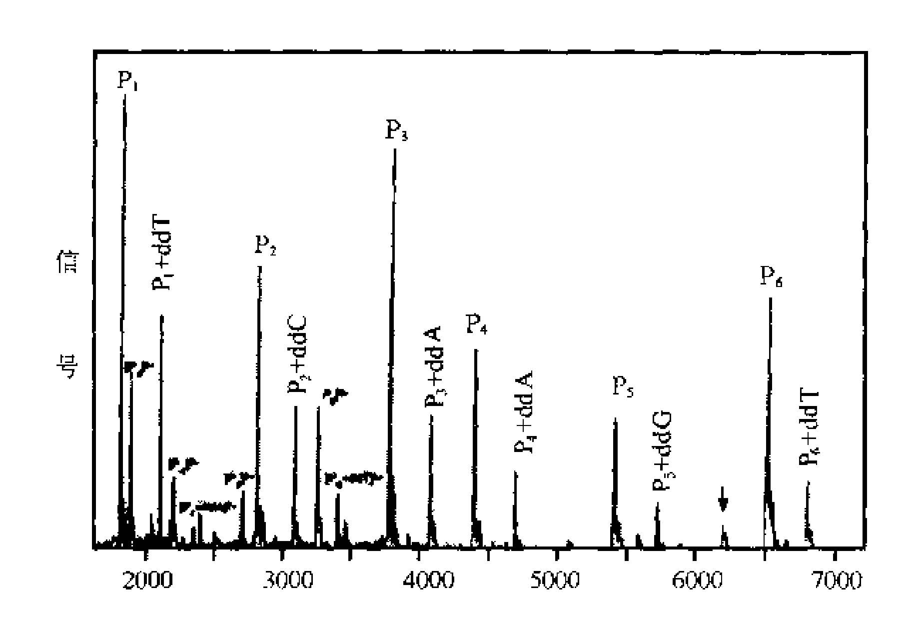


图 16－7 在人类控制区 1021bp 6 个 SNP 标记复合检测质量谱

表 16－4 复合检测 K562 细胞株的人类线粒体 DNA 的 SNP 标记

SNP 标记（mtDNA 位置）	期望碱基（野生型列在前面）	引物质量（Da）	延伸质量（Da）	质量差异（Da）	计算 SNP 碱基
L00146	T 或 C	P1 = 1819.7	2109.8	290.1	T
H00247	C 或 T	P1 = 2813.7	3087.2	273.5	C
H00152	A 或 G	P1 = 3770.0	4068.6	298.6	A
H16311	A 或 G	P1 = 4387.9	4686.8	298.9	A
H16069	G 或 A	P1 = 5411.3	5725.9	314.6	G
L00195	T 或 C	P1 = 6509.4	6798.4	289.0	A

在 SNP 引物中插入不同酶切点比加多聚 T 尾巴（poly－T tail）更适合于利用质量或电泳迁移率来检测多个 SNP。因为多聚 T 方法使 SNP 引物向较高质量方向漂移，延伸产物的质量相对增加，将增加对差异较小杂合子的分离难度。

基本实验分析方法

（1）PCR 扩增：在 20μl PCR 反应体系中，加含有 1μmol/L 引物，5～20ng DNA 模板，1u Ampli Taq GOLD DNA 聚合酶，1×PCR 缓冲液，1.5～2.5mmol/L $MgCl_2$，和 200μmol/L

dNTPs。95℃10min，然后 94℃10s，55℃30s，72℃30s，30 个循环，最后 72℃5min。

（2）PCR 产物用虾 - 碱性磷酸酶（shrimp - alkine phosphatase，SAP）处理，水解未掺入的核苷酸底物 dNTPs。一般在 20μl PCR 产物中加入 1u SAP，在 37℃保温 60min，然后 75℃15min。

（3）SNP 延伸反应：在 20μl PCR 反应体系中，加 5μl SAP 处理的 PCR 产物，1×Taq FS 缓冲液，1.2～2.4u TaqFS，12.5μmol/L ddNTP 混合液，0.5μmol/L 生物素标记 SNP 引物。95℃1min，然后 94℃10s，45～60℃10s（具体温度取决于 SNP 引物的退火温度），70℃10s，25～35 个循环。

（4）MS 分析：通过固相吸附过程，从链亲和素包埋磁珠释放出纯化过的 DNA。含有引物和延伸产物的 DNA 样品与 1～2μl 3 - 羟基甲基吡啶酸基质（3 - hydroxypicolinic acid）溶液混合，点成点。用紫外激光脉冲的 TOF 质谱分析仪分析。用 4 - 聚核苷酸（1272Da）和 15 聚核苷酸（4507Da）二个寡核苷酸质量标准进行质量校正，测定质量，进行分型。

16.2.3 熔解温度曲线法

1. 原理

熔解温度曲线法的原理是不同长度或不同碱基序列的 DNA 片段的熔解温度（Tm）值不同，根据荧光强度随温度变化的熔解温度曲线可以进行基因分型。将双链 DNA 染料 SYBR Green I 加到 PCR 产物中，嵌合到 DNA 双链，在激发光的作用下发出荧光信号。当 PCR 产物在加热过程中渐渐达到 DNA 片段的熔解温度 Tm 时，DNA 双链分开，引起荧光信号的丢失。荧光信号丢失情况代表着每一个 DNA 片段的熔解温度变化情况，而熔解温度变化以熔解曲线的形式表现出来，借助 DASH 仪器（ThermoHybrid US，Franklin，MS）来监测熔解曲线的变化，分析 DNA 样品的基因型。SNP，由于单个碱基的突变，产生或丢失限制性内切酶的切点，选用适当的内切酶酶切 PCR 产物，有酶切点的等位基因片段被切成较短的二个片段，而没有切点的等位基因长度仍不变，导致它们的熔解曲线不同，可以确定出基因型。当然也可以直接利用碱基序列差异引起的 Tm 不同，进行熔解温度曲线分析，只是因单个碱基差异造成的 Tm 差异不是十分大，在熔解曲线上峰不是十分清晰。因此一般选用限制性内切酶酶切 PCR 扩增产物，使之产生不同长度的 DNA 片段，根据熔解温度曲线中峰的位置和数目确定基因型。

2. 用熔解曲线法分析 Y - SNPs ：M170 和 M9

Y - SNP M170 基因座的多态性碱基为 G - T，G 型存在一个 Mn Ⅱ 酶切点，M9 基因座的 G 型存在一个 Hinf I 酶切点，用 PCR 扩增含有多态性碱基的序列，然后用内切酶消化 PCR 产物，酶切产物进行熔解温度曲线的分析。根据熔解温度曲线中出现的峰位置与数目确定基因型。

（1）引物序列

M170 引物 5′ - TGT TTT TTC ATA TTC TGT GCA TT - 3′
5′ - GAC ACA ACC CAC ACT GAA AAA CA - 3′

M9 引物 5′ - CAA AGA AAC GGC CTA AG - 3′
5′ - GAC ATT GAA CGT TTG AAC A - 3′

（2）PCR 反应 在 25μl 反应体系中，含有 2.5μl 10×PCR 缓冲液，0.1mmol/L dNTPs，

2.0mmol/L $MgCl_2$，0.4μmol/L 引物和 2.5u Taq DNA 聚合酶。

（3）扩增循环参数　95℃变性 5min 后，进入循环步骤：94℃30s，60℃45s，72℃45s，循环 30 次，最后 72℃延伸 5min。

（4）内切酶消化　取扩增产物 20μl，M170 基因座用 5u MnⅡ在 37℃消化 18～24h。对于 M9 基因座，用 5u HinfⅠ在 37℃消化 18～24h。

（5）熔解曲线分析　取 10～20μl 酶解产物到另一反应板中，加入检测缓冲液（水、甲酰胺和 SYBR GreenⅠ）使总体积为 50μl。由 McSNP 软件进行分析，以每秒 0.04℃速度缓慢地将温度从 45℃上升到 95℃，观察熔解曲线，总的分析时间为 30min，最后由仪器自动给出 DNA 样品基因型。在仪器上一次可检测 96 个样品。

（6）基因分型　在 M9 和 M170 基因座的熔解温度曲线中，根据等位基因峰位置与数目确定基因分型（图 16－8）。

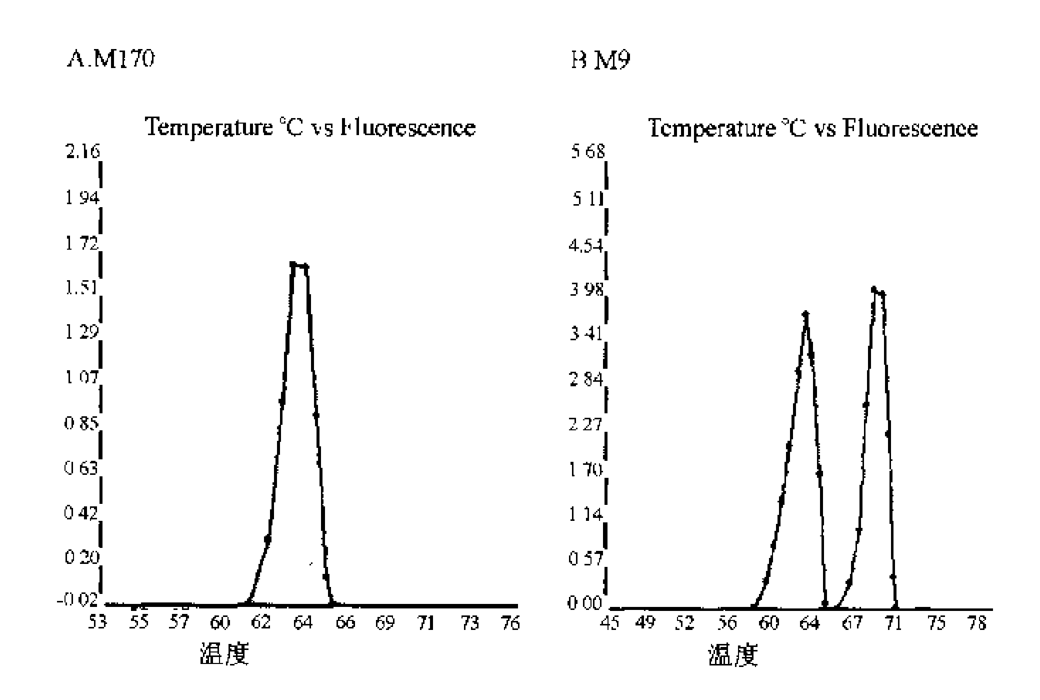


图 16－8　M170 和 M9 基因座的熔解曲线　A：M170；B：M9

为了获得清晰的峰，SNP 分析的 DNA 片段 Tm 值应相差 3℃以上。对于检测 Tm 值在 2～3℃之间的 PCR 产物，应将温度上升速度由 0.04℃/s 调至 0.02℃/s，才能得到好的结果。另外为保证 DNA 片段 Tm 值落在仪器要求的温度范围（35～95℃）内，需要加入甲酰胺以降低 Tm 值。PCR 产物片段的大小也是一个关键因素，PCR 产物片段大小在 50～150bp 范围，可以获得满意的结果。

16.2.4 等位基因特异扩增结合熔解曲线分析法

1. 原理

这种方法是利用等位基因特异引物扩增 SNP 的等位基因，两个上流引物的 3′末端碱基分别对应于 SNP 的两个等位基因，只能扩增其中一个等位基因，反向引物可以扩增两个等位基因。为了提高扩增产物的识别能力，设计反向引物的序列时尽可能使扩增产物片

段最短，在一个等位基因特异引物的5′端连接26个GC尾巴，使其Tm与另一个等位基因的差别明显。PCR扩增结束后，反应物慢慢地升温到95℃，升温的速度为1℃/min（图16-9）。记录SYBR Green Ⅰ的荧光强度减弱情况，得到熔解温度曲线，确定等位基因。

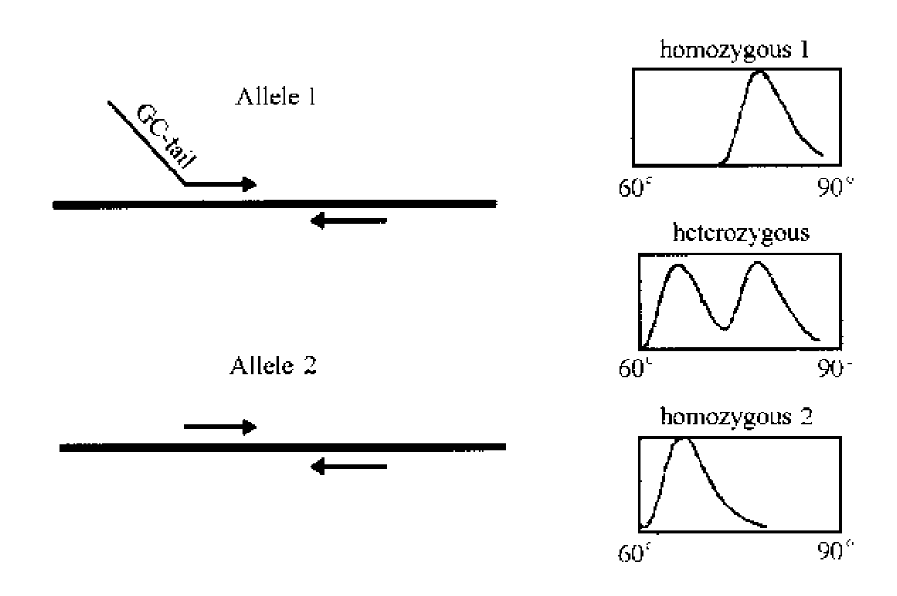


图16-9　Tm-shift基因分型的原理图

在复合反应体系中含有2个等位基因特异引物和一个共用引物，根据模板中存在的SNP等位基因，两个等位基因中的一个被扩增或二个同时扩增出来。纯合子样品，用含有GC尾巴的引物扩增，得到的熔解曲线中的Tm峰在温度范围的最高端，用较短的另一个引物扩增出另一个等位基因，熔解曲线中的Tm峰在较低温度范围。杂合子个体都能被两个正向链引物扩增，熔解曲线拥有两个峰。

2. 方法

(1) PCR扩增　以PON为例说明

引物序列：　5′-TAT TTT CTT GAC CCC TAC TTA CA-3′（短等位基因特异引物）

5′-GCG GGC AGG GCG GCG GGG GCG GGG CC TTT CTT GAC CCC TAC TTA CG-3′（GC尾巴等位基因特异引物）

5′-CCA CGC TAA ACC CAA ATA CAT CTC-3′（共用的反向链引物）

PCR反应：100μl反应体系中，含有40ng模板DNA，0.2μmol/L每种引物，10u Stoffel聚合酶，1×Stoffel缓冲液（10mmol/L KCl，10mmol/L Tris-HCl，pH8.3），另加30mmol/L KCl使之浓度为40mmol/L，2mmol/L $MgCl_2$，50μmol/L dNTPs，75μmol/L dUTP，2u UNG，0.2×SYBR Green I，2.5%甘油。

循环：50℃保温2min，使UND作用，除去因PCR产物引起的污染，然后95℃1min变性，使UND失活。进行二个步骤的37个循环：95℃20s，60℃20s。

(2) 熔解温度　从56℃开始，以每1min增加1℃的变化通过40个温度增加过程得到熔解曲线。SYBR Green I选择亲和双链DNA，当结合到dsDNA时，荧光强度迅速增加，当温度升高，DNA变性，荧光信号即减弱，测定每步温度升高过程中的荧光信号。

(3) 数据分析　熔解曲线是代表每个样品随温度增加荧光变化率曲线（图16-10）。为了获得该曲线，首先，要将测量值转化成相对荧光值：每个温度下的荧光值除以起始

(56℃) 荧光值。相对荧光值可以补偿不同样品槽之间的差异。减去基线，可以大幅度地消除只随温度变化引起的荧光强度变化。基线值来自于每个样品的在 56~65℃范围内的前 9 个数据点的线形回归。根据 Savitzky - Golay 一级导数和数据的修正获得熔解曲线。Savitzky - Golay 值的计算公式：

$$2\times[(X+1)-(X-1)+(X+2)-(X-2)]/10\triangle T$$

式中：(X+1)，(X-1)，(X+2) 和 (X-2) 代表测定点后一个点、前一个点、后二个点和前二个点的荧光值。

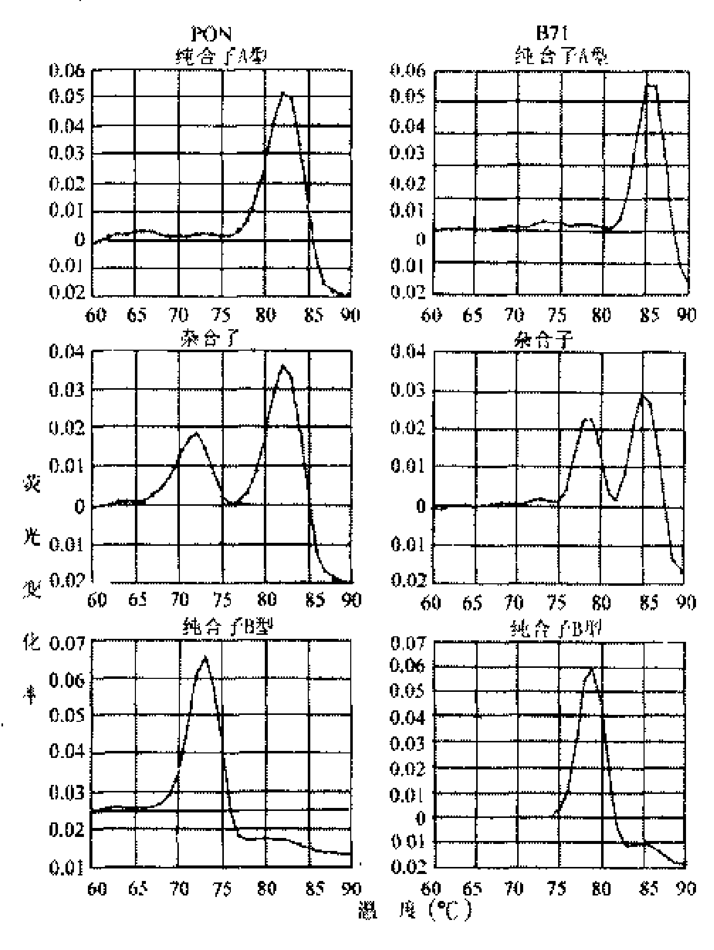


图 16-10 熔解曲线

用该方法分析 100 个样品 SNP 基因座 PON 的基因型，对每个样品在两个相对温度范围内的峰分布作图，进行散点图分析，结果如图 16-11。很清楚地看到每个样品落在三个簇中的一个，图谱证明分析得到的基因型正确、肯定，每一个对应于相应的基因型。

为了更进一步证实分析的有效性，作者将 63 个样品的 Tm-Shift 法得到的分型结果 (PON) 与用特异探针滤条杂交得到的结果比较，二者之间只有一个样品分型不同（表 16-5）。对于直接测序这个样品，确认其分型为 AB 型（杂合子），与 Tm-Shift 得到的结果一致。证实，Tm-Shift 方法分析结果比较准确、可靠。

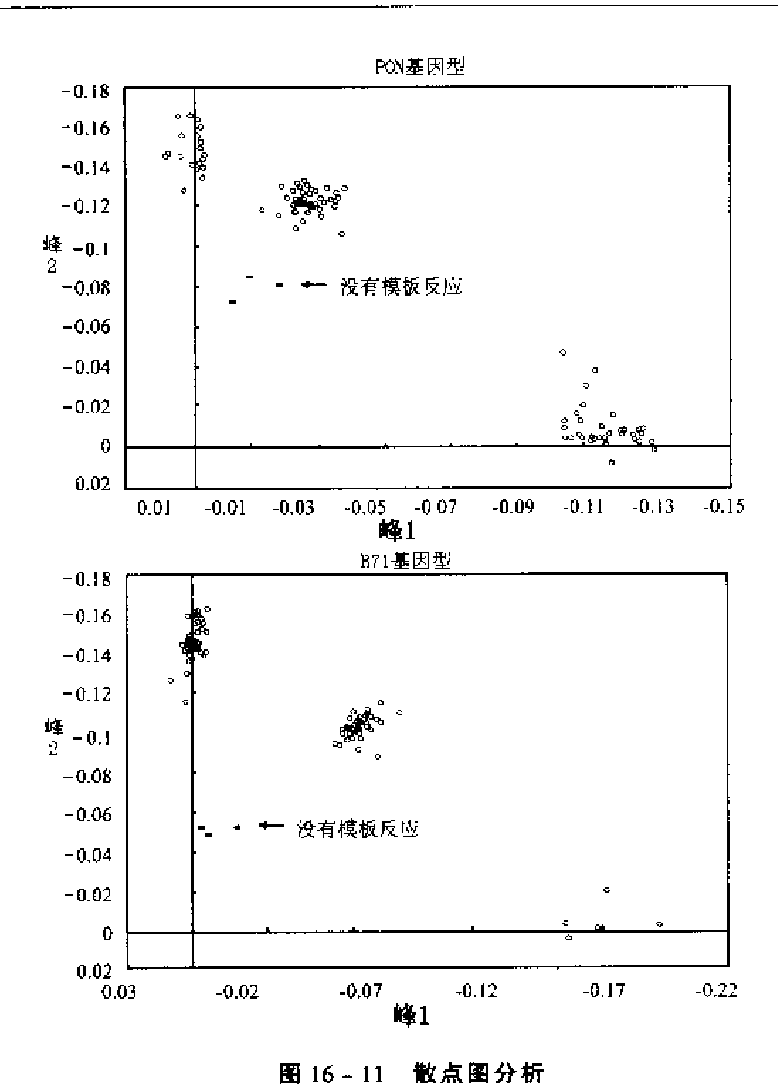


图 16-11　散点图分析

从每个样品的熔解曲线取峰 1 和峰 2 的面积。（○）基因组 DNA；（■）非模板对照

表 16-5　Tm-Shift 与特异探针滤条杂交在 PON 基因座的结果比较

	样品数	Tm-Shift 与特异探针分型一致性
基因型 A	31	31/31
基因型 AB	25	24/25
基因型 B	7	7/7
总数 P	63	62/63

与以荧光探针为基础的方法比较，Tm-Shift 方法所需检材少，一旦等位基因特异引物设计好后，可以很快分析大量样品，40min 的退火和扩增及数据处理，在 GeneAmp 5700 仪上一天可以分析 192 个样品。而且 Tm-Shift 分型法自动化程度高。

16.2.5 动态等位基因特异杂交法

动态等位基因特异杂交（dynamic allele-specific hybridization，DASH）的关键是动力学加热和 DNA 变性一次检测。不需要额外的酶和反应步骤。

图 16 - 12 是 DASH 的原理示意图。用生物标记的引物进行目的 DNA 扩增，使目的片段带有生物素，生物素的产物被链霉亲和素包埋的微孔滴定板固定，不带生物素标记的链用碱洗掉。加一等位基因特异的寡核苷酸探针，在低温度下杂交到目的序列中，形成双链 DNA 区域，双链特异染料结合到该区。插入染料发出的荧光强度与双链 DNA（探针 - 靶 DNA 二倍体）的量有关。然后不断给样品加热升温，不断记录荧光信号的减弱情况。在探针 - 靶 DNA 二倍体的变性（熔解）温度，荧光强度迅速下降。采用一定的缓冲液和染料，可以很容易检测到单个碱基不匹配导致的低熔解温度（Tm）。

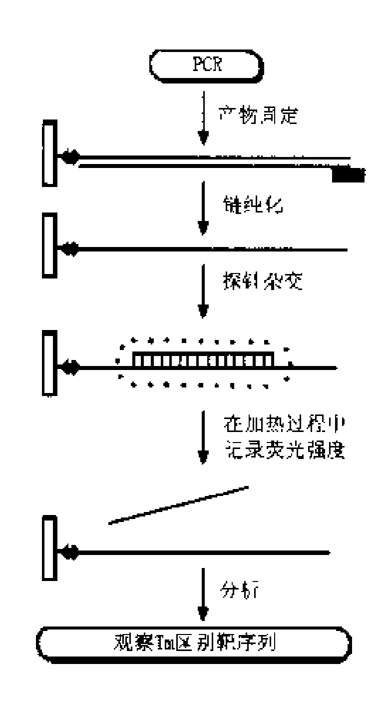


图 16 - 12　DASH 分析原理示意图

16.2.6 LightCycler 结合熔解温度曲线方法

1. 原理

LightCycler 是 Roche 公司开发的一种定量 PCR 仪，不过，它利用二个探针，可以检测 SNP。3'端带有荧光团供体，与靶序列中非多态序列杂交，起固定作用，称之为固定探针，5'端标有 LightCycle Red 640 或 LightCycle Red 705 的另一个探针跨越 SNP 突变位置，称之为敏感探针或突变探针。一旦二个探针与靶序列杂交，它们就紧挨在一起，两个荧光团发生荧光共振能量传递（fluorescence resonanle energy transfer，FRET），在 FRET 过程中，荧光素供体荧光团被 LightCycle 仪器激发，激发的部分能量传递给受体荧光团 LC - Red 640 或 LC - Red 705 荧光。然后受体荧光团发出较长波长的荧光，被光敏二极管测定、记录。这种检测方法可以在 PCR 扩增过程中检测。

如果样品基因型存在突变，敏感探针与靶序列不完全互补，有一个碱基的错配，形成的杂交体不稳定，Tm 相对较低。如果没有突变存在，序列完全互补，杂交有一个较高的 Tm。PCR 反应后，升高温度，观察荧光随温度减弱的曲线，测定敏感探针的 Tm 值可以确定多态性等位基因。图 16 - 13 是其原理示意图。为了获得最佳能量传递，二个探针之间

的空间距离应不大于4bp，因为较大的距离将会导致荧光信号的减少。为了允许最长杂交时间和延缓聚合酶引起的探针位移，每个SNP的二个杂交探针与延伸引物的物理距离越远越好，它们的Tm比延伸引物Tm高5～10℃。为了避免二个探针被Taq聚合酶延伸，探针3′端被磷酸化。

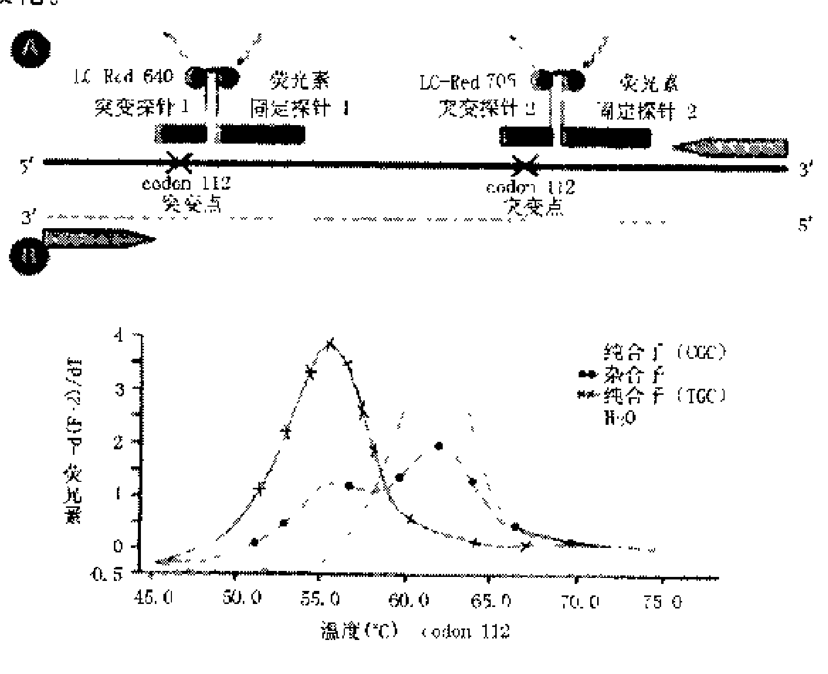


图16－13 lightcycler分析SNP的原理示意图

A：DNA片段、PCR引物、固定探针与突变探针的示意图

B：熔解曲线分析检测SNP（ASPOE基因）codon112的不同基因型，曲线中表明完全互补的纯合子（CCC）的Tm比有一个碱基不匹配的纯合子（TGC）的高。杂合子个体含有二个Tm。

2．快速循环PCR和熔解温度曲线分析sY81，M9和SRY－1532

PCR和熔解温度曲线分析在配有32个样品槽的快速荧光热循环仪中进行。每个SNP在20μl反应体系中进行PCR反应，含有0.25μmol/L每种引物，0．1μmol/L的固定探针和敏感探针，2μl的10×缓冲液（含有Taq DNA聚合酶，Taq PCR缓冲液，dNTP，10mmol/L $MgCl_2$）。每个SNP基因座的引物与探针序列如表16－6和表16－7。

表16－6 SNP扩增引物序列

SNP	大小（bp）	正向链引物	反向链引物
sY81	209	5′AGG CAC TGG TCA CAA TCAAG3′	5′AAT GGA AAA TAC AGC TCC CC3′
M9	341	5′GCA GCA TAT AAA ACT TTC AGG3′	5′AAA ACC TAA CTT TGC TCA AGC 3′
SRY－1532	165	5′TCC TTA GCA ACC ATT AAT CTG G3′	5′ATA GCA AAA AAT GAC ACA AGG CA3′

表 16－7 序列特异杂交探针

SNP	GBD no.	杂交探针序列
sY81	S76940	AAA ATG TAG GTT TTA TTA TTA TAT TTC ATT GT－X（固定）
		LC－Red 640－AAC AAA AGT CCG TGA GAT－P（sensor）
M9	G42825	CAG AAC TGC AAA GAA ACG GCC TAA－X（anchor）
		LC－Red 640－TGG TTG AAT GCT CTT TAT TTT TCT T－P（sensor）
SRY－1532	L08063	ATC ATT CAG TAT CTG GCC TCT TGT ATC T－X（anchor）
		LC－Red 705－CTT TTT CAC ACA GTG TAA CAT TTT CAA－P（sensor）

X 代表荧光素，P 为 3′端磷酸；带下划线的核苷酸是每个 SNP 分型涉及的碱基

LightCycler 方法检测特异序列和探针杂交分析包括 4 个循环过程：(1) 模板 DNA 的起始变性；(2) 靶 DNA 的扩增；(3) 熔解温度曲线分析；(4) 仪器的冷却。具体的 SNP 分析实验条件如下：

过程	循环数	步骤 1	步骤 2	步骤 3
变性	1	95℃30s		
循环扩增		95℃0s	sY81：60℃10s	72℃8s
			M9：59℃10s	72℃13s
			SRY－1532：60℃10s	72℃7s
熔解	1	95℃0s	40℃60s	85℃0s
冷却	1	40℃0s		

步骤 1 和 2 的温度转换速度为 20℃/s，步骤 3 的温度转换速度为 0.2℃/s。

SNP 的熔解温度如表 16－8 所示。标有同一荧光基团的 SNP 的 Tm 不同，彼此不重叠。图 16－14 为 sY18 位点的分析结果。该仪器可以检测三色荧光，不同 SNP 可以复合在一起检测。

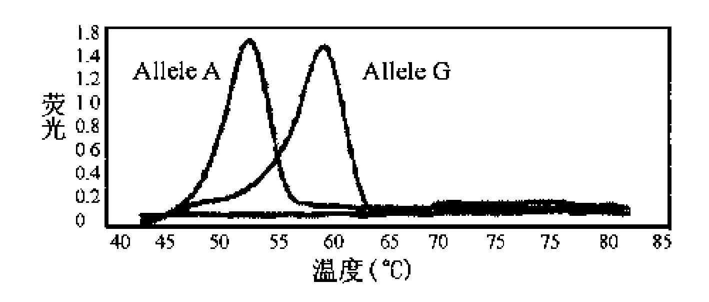


图 16－14 sY81 的二个等位基因的熔解温度曲线分析结果

表 16-8　三个分析的 SNP 的熔解温度

SNP	平均熔解温度		不匹配	标记
	等位基因 0	等位基因 1		
sY81	51.35℃	57.48℃	G-T	LC-Red640
M9	57.85℃	61.78℃	G-C	LC-Red640
SRY-1532	59.26℃	64.02℃	G-T	LC-Red705

16.2.7 分子信号法

分子信号（信标）是发卡状分子杂交探针，用来检测 PCR 荧光产物形成。分子信号过程非常简单，分子信号加到 PCR 反应混合液中，进行 PCR 扩增，产物积累，分子信号结合到它们特异的靶序列上，每个分子信号-靶序列杂交体发出特定分子信号的一定波长的荧光，发出的荧光被荧光仪实时记录下来或在扩增结束后检测。

分子信号含有一个荧光分子，它的另一端有一个非荧光淬灭基团（图 16-15）。寡核苷酸的两端序列是互补的，形成一个手臂（或叫茎）结构，插入的环（探针区域）的序列与扩增产物内部的序列互补。当分子信号游离在溶液时，它成为发卡结构，荧光分子与荧光淬灭基团在空间距离上挨在一起，十分接近，使荧光分子被淬灭。反过来，如果溶液中存在分子信号的互补序列，分子信号自发地结合到合成的寡核苷酸靶序列上（图 16-16），探针分子构象发生变化，荧光分子与淬灭基团在空间上相隔甚远，致使荧光分子不能被淬灭，分子信号发出荧光。某一循环或循环后的荧光量取决于那时形成的特异的 PCR 产物。同一管内可以使用不同的荧光素，一旦分子信号结合上去，可以同时检测到同一 PCR 反应中积累的多个靶序列。分子信号分析方法省去了凝胶电泳步骤，也减少了潜在的污染。

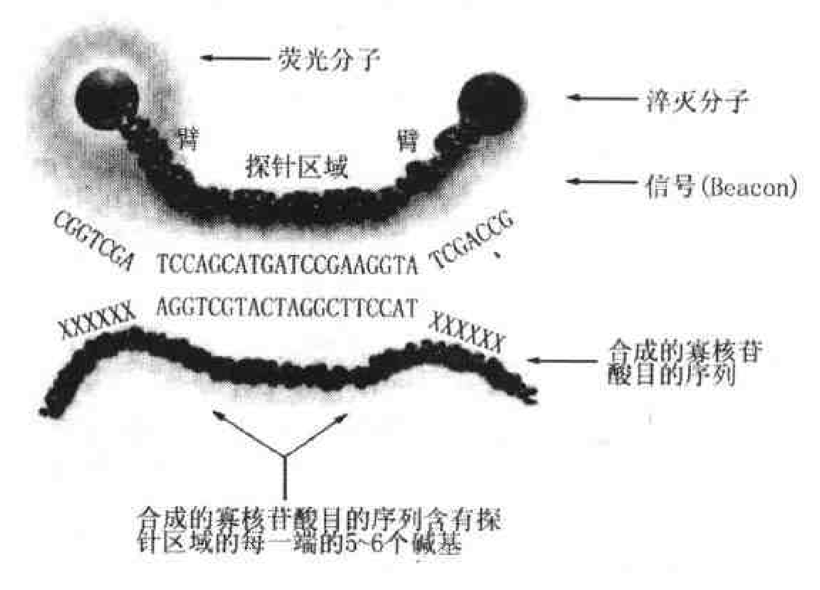


图 16-15　分子信号结构

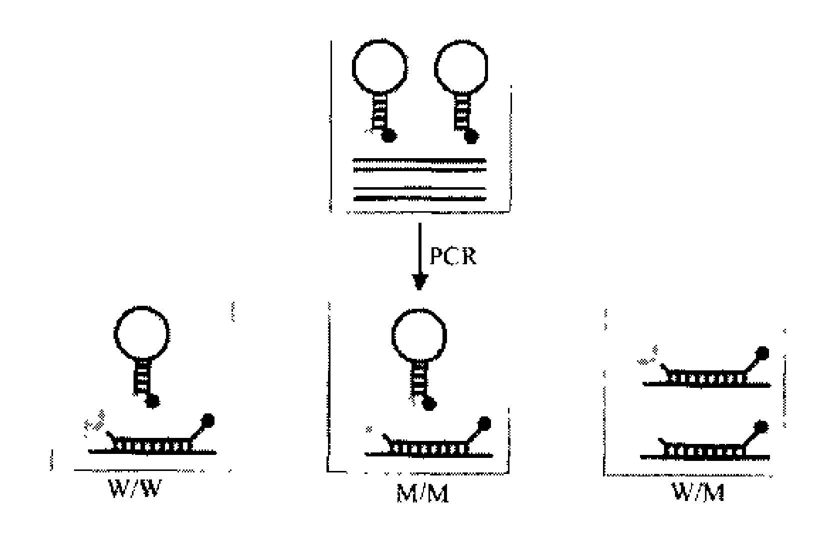


图 16－16 分析信号检测 SNP 原理图

2. 分子信号的设计

为了成功检测 PCR 反应，要求在 PCR 退火温度分子信号与靶序列杂交，游离的分子信号紧密地在一起，在退火温度没有荧光。一般，分子信号环的退火温度要比 PCR 引物退火温度高 7～10℃，保证在 PCR 反应退火时分子信号－靶序列不结合。一旦探针序列确定下来，加上二个互补的茎序列，探针的两端各一个。茎序列的长度、碱基和 GC 含量序列必须满足茎与它的互补序列在高于 PCR 引物退火温度 7～10℃的温度下分离。在 PCR 引物退火温度下，未杂交的信号还保持发卡结构，没有荧光。

根据 GC 原则，可以算出探针与靶序列杂交的退火温度。一般，探针序列长度为 15～33 个核苷酸。游离分子的构象设计成发卡结构，过长的茎会降低分子信号与靶序列的杂交速度。

分子信号序列设计、合成后，调整 PCR 缓冲液条件，确定分子信号退火特性，以保证分子信号信能良好，符合预期的要求。图 16－17 为典型分子信号的三个退火曲线的退火图，一条是单独分子信号的曲线，第二条是分子信号完全互补单链寡核苷酸靶序列，第三条是分子信号与靶序列有一个碱基不匹配的曲线。

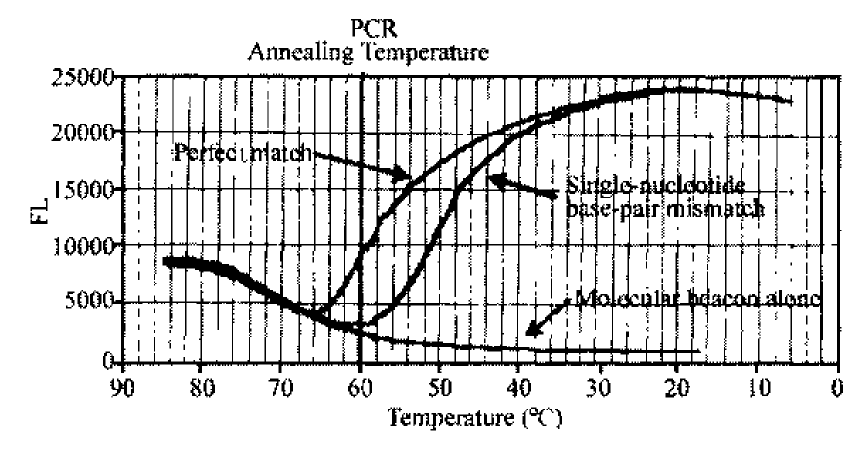


图 16－17 分子信号的熔解温度曲线

16.2.8 TaqMan 法

TaqMan 技术是 PE 公司开发的，目前已开始应用于基因检测。该技术的基本原理如图 16-18 所示。它利用了 Taq 酶的 5′外切酶活性，合成一个能与 PCR 产物杂交的探针，该探针的 5′端标记一个荧光分子，3′端标记另一个荧光分子。其中 3′端荧光分子能够吸收 5′端荧光分子发出的荧光，因此正常情况下该探针检测不到 5′端荧光分子发出的荧光，只能检测到 3′端荧光分子的荧光信号，但当溶液中有 PCR 产物（模板）时，该探针与模板结合，激发 Taq 酶的 3′外切酶活性，将探针 5′端连接的荧光分子从探针上切割下来从而发出荧光，切割的荧光分子数与 PCR 产物的数量成比例。设计等位基因特异的引物，进行等位基因特异扩增，检测 PCR 反应液中的荧光强度确定 SNP 分型。其主要不足是：（1）采用荧光淬灭及双末端标记技术，因此淬火难以彻底，本底较高；（2）采用酶外切活性，因此定量时受酶性能影响；（3）探针标记成本较高，不便普及应用。TaqMan 技术在定量分析及基因遗传突变分析等领域有广泛应用。

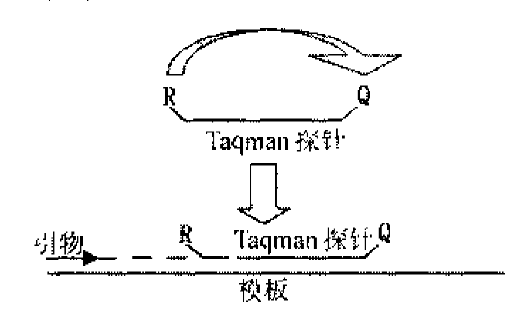


图 16-18　Taqman 原理图

16.2.9 焦磷酸测序法

焦磷酸测序法原理是利用一个级联化学反应监测核酸链延伸反应，同时产生一个焦磷酸，焦磷酸在 Sulfrylase 催化下，与 APS 反应生成 ATP。ATP 在荧光素酶的作用下，发出的荧光由 PSQ96 系统检测。检测到的发光强度与掺入核苷酸数量成正比（图 16-19）。

利用焦磷酸测序在 10min 内可以检测 96 个样品，具有快速、高通量，准确率高等特点，不需要制备胶，不需要毛细管，不需要荧光染料和同位素，也不需要特异 DNA 标记。

16.2.10 其他分析方法

除了上述介绍的方法外，还有以杂交为原理的 DNA 杂交芯片技术分析 SNP，将所有的序列特异的寡核苷酸探针经过处理点样在尼龙膜/玻璃载片或硅片等固相支持物上，与待检测的样品 DNA 进行杂交，根据检测到的信号位置进行基因分型，PM 系统就是一个很小的芯片，目前已有人开发出了拥有 20 000 多个探针的线粒体 DNA 分析芯片。另外，在线粒体 DNA 章节介绍的微型测序方法也可以进行 SNP 分析，具体参见线粒体章节。

16.3 法医学应用的 SNP 标记

SNP 广泛存在于基因组中，理论上每 300～1 000bp 发现一个 SNP，是目前数量最多的遗传标记，被认为是第三代遗传标记，也是未来广泛应用的标记之一。事实上，已有许多 SNP 标记广泛应用于法医学领域，进行个体识别与亲权鉴定。如线粒体 DNA 的 D-环区的序列、白细胞抗原 HLA Ⅰ类和 Ⅱ类标记、ABO 和 MN 等血清型标记、蛋白酶型，有 5 个

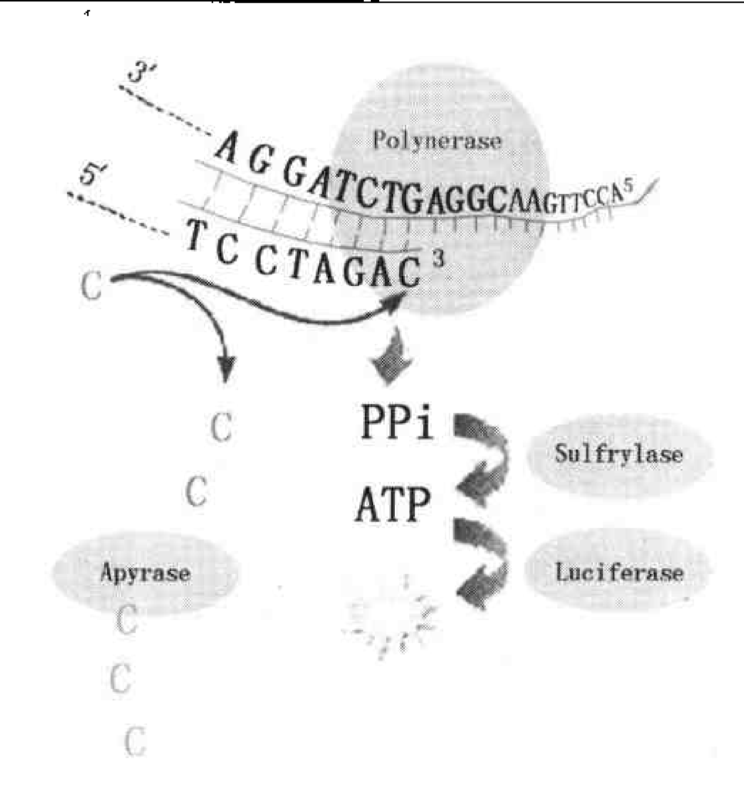


图 16－19 焦磷酸测序法分析 SNP 原理示意图

标记组成的 Polymaker（PM）系统，等等，都是 SNP 标记。

16.3.1 HLA 标记

HLA 系统的遗传基因定位于第 6 号染色体短臂 2 区 1 带 3 亚区，这个区域称为 HLA 区或主要组织相容性复合体（major histocompatibility complex，MHC）。控制 β2 微球蛋白的基因在第 15 号染色体上。按其编码产物的结构、表达方式、组织分布与功能可将 HLA 基因区分为三类。

I 类基因区域包括 HLA－A，B，C，E，F，G，H 和 J 8 个座位。HLA－A，B 和 C 基因分别编码 A、B 和 C 抗原的 a 链，其中 A 有 50 个等位基因，B 有 97 个，C 有 34 个。E，F 和 G 座位的产物功能不明，H 和 J 是假基因。

II 类基因又称 D 区，主要包括 HLA－DR，DP 和 DQ 三个亚区，每个亚区均有 A 基因和 B 基因，分别编码 a 链和 β 链。DRA 基因几乎无多态性，只有 2 个等位基因。DRB1 拥有 106 个等位基因，B3 有 4 个，B4 有 5 个，B5 有 5 个。DQ，DP 各有 2 对 A 基因、B 基因。DQA 1 有 15 个等位基因，B1 有 26 个。DPA1 有 8 个，B1 有 59 个。DRB1，B3，B4，B5，DQA1，B1，DPA1，B1 都是功能基因，DRB2，DPA2，B2 是假基因，DQA2，B2 的基因产物未能检出。

HLA 研究进展迅速，先后证实了 HLA－A，B，C，D，DR，DP，DQ 七个座位，并研究了白细胞配型与皮肤、肾移植的关系，HLA 与疾病的关联性以及 HLA 的群体遗传学等。

HLA 多态性分析主要是用以 PCR 为基础的方法，如 PCR－RFLP，PCR－SSP，PCR－SSOP 和 SSCP 方法，其中 PCR－SSOP 方法研究最早、应用最广。HLA 分析芯片正在研究中，在不久的将来面世。

16.3.2 Polymarker（AmpliType PM + DQA1）分析系统

Polymarker（AmpliType PM + DQA1）分析系统可以同时扩增 6 个遗传标记［HLA DQA1，LDLR（low－density lipoprotein receptor，低密度脂蛋白受体），GYPA（glycophorin A，血型糖蛋白 A），HBGG（hemoglobin G`gammsglobin，血红蛋白 Gγ 球蛋白），D7S8，Gc（group－specific component，组特异成分）］。表 16－9 为其染色体定位。所有 6 个基因座可以用反向等位基因特异寡核苷酸探针杂交检测。5 个 Polymarker 基因座共检出 972 个基因型，HLA DQA1 基因型共检出 28 个基因座，PM + DQA1 系统共有 27，216 个基因型。表 16－10，表 16－11 为四个人种的等位基因频率。

表 16－9　PM 系统组成的有关基因座

	HLA DQA1	LDLR	GYPA	HBGG	D7S8	GC
染色体定位	6p21.3	19p13.1 － 13.3	4q28 － 31	11p15.5	7q22 － 31.1	4q11 － 13
PCR 产物片段大小（bp）	239/242	214	190	172	151	138
等位基因数	7[a]	2	2[b]	3	2	3

a：HLADQA1 4.2 和 4.3 型等位基因可以检出，但在此分析中彼此没有差别。

b：可以检出 GYPA A 和 A 引物等位基因，但此系统中彼此没有差异。

表 16－10　PM 系统 4 个人种的等位基因的频率

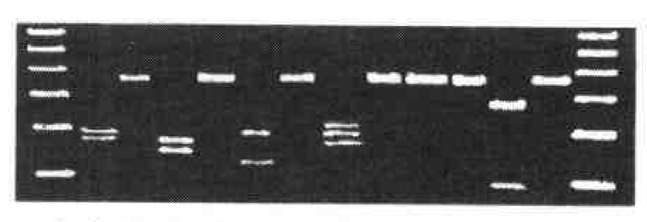
等位基因	非－美人群	高家索人群	SE 修斯敦人群	SW 修斯敦人群
LDLR A	0.224	0.453	0.415	0.563
LDLR B	0.776	0.547	0.585	0.438
GYPA A	0.476	0.584	0.532	0.656
GYPA B	0.521	0.416	0.468	0.344
HBGG A	0.507	0.470	0.426	0.344
HBGG B	0.197	0.524	0.548	0.609
HGBB C	0.297	0.007	0.027	0.047
D7S8 A	0.614	0.615	0.585	0.682
D7S8 B	0.386	0.385	0.415	0.318
Gc A	0.1023	0.257	0.277	0.271
Gc B	0.707	0.172	0.223	0.208
Gc C	0.190	0.571	0.500	0.521

表 16－11　HLA DQA1 基因座的等位基因频率

等位基因	非－美人群	高家索人群	修斯敦人群
1.1	0.125	0.158	0.105
1.2	0.329	0.190	0.130
1.3	0.058	0.073	0.053
2	0.130	0.145	0.115
3	0.090	0.192	0.218
4.1	0.185	0.214	0.269
4.2/4.3	0.083	0.028	0.110

第十七章　种属与性别鉴定

17.1 种属鉴定



2. 方法 2

引物序列 5′ - CAT CGA CCT TCC AGC CCC ATC AAA CAT - 3′

5′ - TGT TCT ACT GGT TGG CCT CCA AT TCA - 3′

此对引物是 Zehner（1998）报道，可以扩增 981bp Cytochrome b 基因片段。

扩增体系 在 50μl 反应体系中，含有 100ng DNA，20pmol 引物，200μmol/L dNTPs，1 × PCR 缓冲液，160μg/ml BSA，0.2u Taq DNA 聚合酶。

PCR 循环参数 94℃3min，94℃ 30s，55℃ 30 s，72℃ 30s，30 个循环后，72℃ 8min，1 个循环。

PCR - RFLP PCR 扩增产物在 1% 琼脂糖凝胶中电泳分离，纯化回收 981bp。如果扩增特异性较好，在 PAGE 中没有非特异带，不需要进行 PCR 扩增产物的纯化，可以直接进行 RFLP 分析。按内切酶生产厂家使用要求进行 Alu Ⅰ 和 Nco Ⅰ 的消化，酶切产物在 6% PAG 电泳。人和各种动物的酶切图谱不同（表 17 - 1），由此可以区分动物与人，不同品种的猪之间也存在着差异，并且真菌 mtDNA 不能扩增出 981bp 片段。

表 17 - 1 Cyto b 基因的 PCR - RFLP 分析结果

内切酶＼物种	牛	猪	羊	狗	猫	野兔	家兔	乌龟	鸡	人
Alu Ⅰ	453	580	517	463	655	450	834	-	-	978
	322	243	450	387	206	374	157			3
	206	131	14	101	120	157				
		27		30						
Nco Ⅰ	-	-	-	-	-	-	- -	623	623	-
								349	349	

17.1.2 28srDNA 分析

28sRNA 大亚基序列进化慢，相当保守，而 28s rDNA 则进化快，物种之间变化较大。Natio（1992）等设计引物扩增 28s rDNA（图 17 - 2），发现人和动物有差异（图 17 - 3b）。人 DNA 的扩增片段为 108bp、104 bp、101 bp 和 99 bp，主要产物为 99 bp 片段。出现 4 个片段的原因可能是 28s rDNA 定位于 5 个不同的染色体上，但不同个体的带是一致的，没有多态性（图 17 - 3a）。设计此区域特异的引物进行 PCR 扩增，进行种属鉴定。

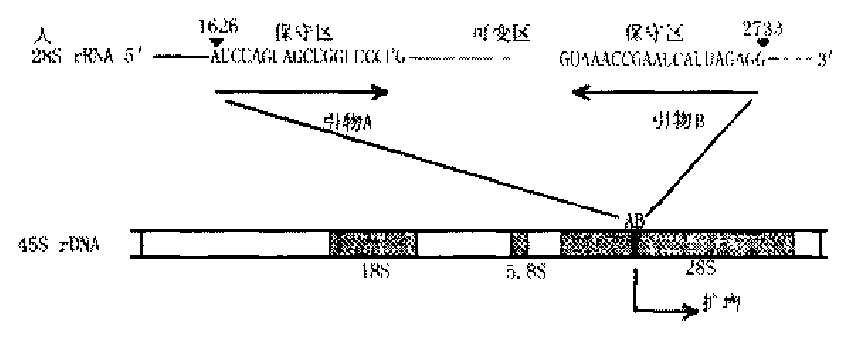


图 17 - 2 引物结合区

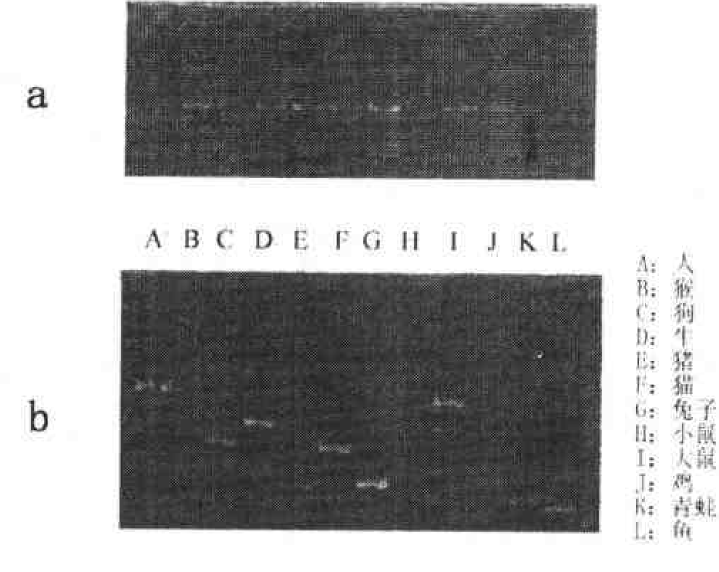


图 17-3 28srDNA 分析物种

a：12 个不同的人；b：不同物种的鉴定

引物序列 5′-ATC TAG TAG CTG GTT CCC TC-3′

5′-CCT CTA ATC ATT CGC TTT AC-3′

PCR 反应 在 30μl 反应体系中，含有 10~100ng DNA，1×PCR 缓冲液（50mmol/L KCl，1.5mmol/L $MgCl_2$，10mmol/L Tris-HCl（pH8.3）），0.2μmol/L 每种引物，200μmol/L dNTPs，1u Taq DNA 聚合酶。

PCR 循环 94℃1min，58℃min，72 ℃1min，循环 30 次。

PCR 扩增产物在 8%聚丙烯酰胺凝胶（19:1）电泳，银染检测。

17.1.3 Alu 序列分析

前面已提到 Alu 序列为哺乳动物特异的序列，有文献报道用 PCR 扩增 Alu 序列进行种属鉴定。如用 Alu9.1 和 9.2 引物进行 PCR 扩增，只有人和猴可得到 130bp 的特异性片段，其他动物得不到扩增产物，由此可以鉴别是否为人 DNA。

引物序列 Alu9.1 5′-GGC ACT TTG GGA GGC CAA GG-3′

Alu9.2 5′-TAC AAG CTT GTG CCA CCA TGC CAA C-3′

在反应体系中，含有 10~100ng DNA，0.2μmol/L 每种引物，200μmol/L dNTPs，1u Taq DNA 聚合酶，50mmol/L KCl，1.5mmol/L $MgCl_2$，10mmol/L Tris-HCl。PCR 反应在 95℃变性 5min 后，进行 94℃30s，55℃30s，72℃1min，30 个循环，最后在 72℃保温 10min。

扩增产物在 2%琼脂糖凝胶中电泳分离，EB 染色，紫外灯下观察结果。

17.1.4 SON 基因 3′端非翻译区（SON 3′UTR）分析

SON 基因定位于人 21 号染色体，用 PCR 分析哺乳动物，发现 PCR 扩增片段大小在 160~175bp 范围，具有扩增片段长度多态性，如人 DNA 片段为 175bp，狗 174bp，猫为 172bp，仓鼠 171bp，猴 168bp，老鼠 160bp（图 17-4）。与人 DNA 比较，存在许多种属特异的 A 或 T 的缺失与插入，物种之间有差异（图 17-5），但物种内没有差异。

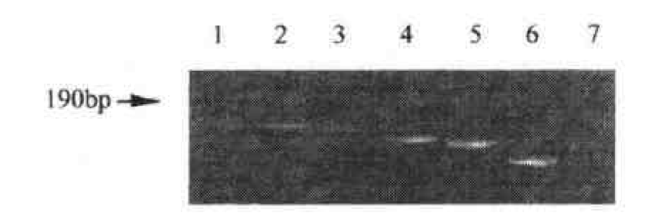


图 17－4　人与数种动物的 SON 基因 3′端非翻译区多态性

1：人；2：狗；3：猫；4：猴；5：硕鼠；6：小鼠；7：阴性对照

```
人    TCAGCTACACCTAGATAGACGTAAAATGATAATTAAAATGCTGTAACCAA
猴子  TCAGCTACACCTAGATAGACGTAAAATGATAATTAAAATGCTGTAACCAA
猫    TCAGCTACACCTAGATAGACGTTAAATGATAATTAAAATGCTGTAACCAA
狗    TCAGCTACACCTAGATAGACGTTAAATGATAATTAAAATGCTGTAACCAA
老鼠  TCAGCTACACCTAGATAGACGTAAAT----AA--AAAT-GCTGTAACCGA
仓鼠  TCAGCTACACCTAGATAGACATAAATTAATAA--AAATCGTCTTAACCAA
      **************************  *  **      **  ***  *   ****  *
人    CTTATCTAATAAAA-TTGGCAACCAGCCACTATTTTGTTGACTATGAGAA
猴子  CTTATCTAATAAAA-TTGGCAACCAGCCACTATTTTGTTGACTATGAGAA
猫    CTTATCTAATAAAA-TCGGCACCCAGCCACTATTTTGTTGACT--GAGAA
狗    CTTATCTAATAAAA-TCGGCAACCAGCCACTATTTTGTTGACTATGAGAA
老鼠  CTTATCTAATAAAAATTGGCACCCAGCCGCTATTTTGTTGACT--GAGAA
仓鼠  CTTATCTAATAAAAATTGGCAACCAGCCACTATTTTGTTGACTATGAGAA
      *****************   ****  ******  *************     ****
人    AGTTAAAAGTTTATGTTAATTTTTAGGGTCTGATAGAATATTTCATGTGT
猴子  AGTTAA-------TGTTAATTTTTAGGGTCCGATAGAATATTTCATGTGT
猫    AGTTTAA-GTTTATGGTAATTTTTAGGGTCTGATAGAATATTCCATGTGT
狗    AGTTAAA-GTTTATGTTAATTTTTAGGGTCTGATAGAATATTTCATGTGT
老鼠  AGTTTAT-GTTAAT-----TTTTTAGGGTCTGATAGA-TATT-CATGTGT
仓鼠  AGTTTAA-GTTTATGTTAATTTTTAGGGTCTGATAGA-TATT-CATGTGT
      ****  *     *       ***********  ******  ****  *******
人    ATTACAGTGGTATTCATATGCTATGT
猴子  ATTCCAGTGGTATTCATATGCTATGT
猫    ATTACAGTGGTATTCATATGCTATGT
狗    ATTACAGTGGTATTCATATGCTATGT
老鼠  ATTACAGTGGTATTCATATGCTATGT
仓鼠  ATTACAGTGGTATTCATATGCTATGT
      ***  ******* ********* ******
```

图 17－5 哺乳动物 SON 3′UTR 序列

下划线部分为 PCR 引物结合区，* 表示序列相同处，－ 表示缺失

方法：

PCR 引物 5′ - ACA TAG CAT ATG AAT ACC - 3′

5′ - GTC TAT CTA GGT GTA GCT CA - 3′

PCR 反应体系 在 100μl 反应体系中，含有 1μmol/L 引物，200μmol/L dNTPs，1.5mmol/L $MgCl_2$，5～50ng 基因组 DNA，2.5u Taq DNA 聚合酶。

PCR 循环参数 95℃1min，54℃1min，72℃2min，循环 30 次。

PCR 产物分析 PCR 产物加在 10%非变性聚丙烯酰胺凝胶中，100V 电压下电泳过夜。

SON 3′UTR 多态性分析方法具有以下优点：

（1）单一 PCR 产物，结果容易解释；

（2）PCR 扩增产物片段小，容易扩增，利于降解样品分析。

此外，1994 年 Lee 等人用随机引物扩增（Random amplified polymorphic DNA，RAPD）方法对多种动物的基因组 DNA 进行扩增，建立多种动物 RAPD 指纹图，通过与已知动物 DNA 比对，可以鉴别出牛、山羊、狗、猪、小鼠、鸡及鸭等动物。只是 RAPD 技术难度大，重复性差些，还需要已知阳性对照。

17.2 性别检验

17.2.1 概述

用 DNA 技术鉴定性别，自 20 世纪 80 年代中期开始各国学者就报道了 Y 染色体特异 DNA 探针杂交以及扩增 Y 染色体特异 DNA 的序列等多种方法检测性别。20 世纪 90 年代开始用 PCR 法扩增 X、Y 两条染色体特异 DNA 片段鉴定性别。如 Amelogenin（Y - 染色体长臂）和睾丸决定因子 TDG 基因（Y - 染色体短臂）。由于 DNA 分析技术鉴定性别具有准确、微量等优点，如今已成为法医物证性别鉴定的主要方法。

要求用于性别鉴定的方法能够同时或平行检验 X、Y 染色体。如果只分析检验 Y 染色体，不同时检验 X 染色体，结果解释比较复杂，因为有以下三种可能情况会导致没有 Y - 特异产物：（1）只有女性 DNA；（2）有男性 DNA，但量太少，不能得到扩增产物；（3）PCR 反应失败。因此在性别检验中，如果没有检出 Y - 特异产物，就难以确定究竟是属于哪种情况，无法判定性别。

还有，有一些遗传标记，虽同时检验了 X 和 Y 染色体，但靶 X 和靶 Y 的特异模板拷贝数不同，如 α - 卫星为基础的 Y 染色体的 alphoid 只有 100 个拷贝，而 X 染色体上的 alphoid 大约为 5 000 拷贝，在这种情况下，男性样品易被认为是女性。除了模板 DNA 拷贝数外，PCR 引物和 PCR 参数也会影响检验结果。

在法医 DNA 分析中具有实际应用价值的性别检验方法是 ZFY/ZFX 基因扩增和 Amelogenin 基因扩增检验。尤其是 Amelogenin 基因扩增，同时检验 X 和 Y 染色体，结果解释简单，X 和 Y 特异谱带之间只有 6bp 差异，即使严重降解的 DNA 也不会出现只扩增二个等位基因中的一个等位基因，而且 Amelogenin 基因可以与 STR 基因座复合扩增。在此主要介绍扩增 ZFY/ZFX 基因鉴定性别和 Amelogenin 基因扩增进行性别鉴定。

17.2.2 ZFY/ZFX 基因

ZFY 基因位于 Y 染色体短臂，它编码一种锌指状蛋白，功能尚不清楚，距离 TDF 基

因的候选基因 SRY 很近。设计三个引物，同时扩增 Y 染色体及 X 染色体特异的 ZFY 基因和 ZFX 基因特异 DNA 片段确定性别，位于 Y－染色体的 ZFY 基因长 340bp，位于 X 染色体的 ZFX 长 488bp。因此女性 DNA 只有一条 488bp 的谱带，男性 DNA 有 340bp 和 488bp 二条谱带。

引物序列：

ZFXY　5′－ATT TGT TCT AAG TGC CAT ATT CTC T－3′

此引物为 ZFX 和 ZFY 基因二者同源序列通用的引物

ZFY_3　5′－CAT CAG CTG AAG CTT GAT GAC ACA CT－3′

此引物为 ZFY 基因特异的 3′端引物

ZFX_3　5′－AGA CAC ACC TAC TGA GCA AAA TGT ATA－3′

此引物为 ZFX 基因特异的 3′端引物

PCR 反应：在 20μl 反应体系中，加入 10mmol/L Tris－HCl（pH8.0），50mmol/L KCl，1.5mmol/L $MgCl_2$，0.1μmol/L ZFXY 引物，各 0.05μmol/L 的 ZFX_3 和 ZFY_3 引物，0.5uTaq DNA 聚合酶。

PCR 循环参数：94℃ 4min 后，按 94℃30s，55℃ 60s，72℃ 75s，循环 35 次。72℃保温 5min。

PCR 扩增产物检测：10～15μl PCR 扩增产物加在 2％琼脂糖凝胶中，DNA/Hind Ⅲ为分子量标准，以 1× TAE 为缓冲液，50～100V 电压下电泳分离 10～15min。电泳后凝胶用 EB 染色检测，紫外灯下观察结果。

17.2.3 Amelogenin 基因

牙釉质蛋白基因（Amelogenin）为 X、Y 染色体共有。Nakahori 等人（1991）对人 Amelogenin 基因进行了测序，AmelX－gene 大小为 2872bp，定位于 X－染色体 p22 区域，AmelY－gene 大小为 3272bp，定位于p12.2 区域。X、Y 染色体的基因缺失序列不同（表 17－2），内含子的长度不同，扩增分析此基因片段长度，可以进行性别检验。Amel X、Y 有 19 个同源区，大小在 22～80bp 范围（表 17－3，图 17－6）。

表 17－2　3272bp AmelY－gene 和 2872bp Amel X－gene 的缺失位置和大小

缺失在 Amel X－gene 的定位（bp）	缺失大小（bp）	缺失在 Amel Y 内含子或外显子的定位	缺失在 Amel Y－gene 的定位（bp）
585－769	185	内含子 2	495
865－868	4	内含子 2	690
951－955	5	内含子 2	772
1761	1	内含子 3	1581
2720－2721	2	外显子 4	2539

表 17－3 Amel X 和 AmelY 同源区

编号	Amel X 上定位（bp）	同源区大小（bp）	Amel Y 定位（bp）
1	186～265	80	186～265
2	287～310	24	287～310
3	368～412	44	375～418
4	899～927	29	721～749
6	1005～1026	22	822～843
7	1342～1364	24	1157～1181
8	1375～1399	24	1192～1215
9	1400～1423	24	1217～1240
10	1425～1462	38	1242～1280
11	1512～1533	33	1330～1362
12	1599～1634	26	1417～1442
13	1655～1680	26	1473～1498
14	1711～1739	29	1532～1560
15	1820～1853	34	1642～1674
16	1965～2020	56	1785～1840
17	2058～2095	38	1878～1915
18	2197～2256	60	2017～2076
19	2398～2441	44	2218～2261

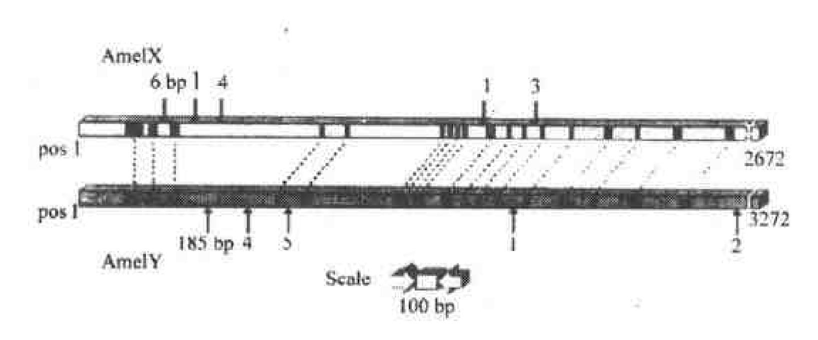


图 17－6 AmelX、AmelY 基因碱基同源及缺失部位、缺失碱基数目示意图

黑色部位为可设计 X、Y 同源序列引物所在部位，箭头所示为同源缺失部位，数字为缺失碱基数。

Nakahori（1991）最早用 PCR 扩增单拷贝的 X、Y 同源区域进行性别测定。随后，Aksne（1991，1992），Bailey（1992），Faerman（1995），Sullivan（1993）等先后设计了多对 PCR 引物进行 Amelogenin 基因检验，扩增片段长度各不相同，最常用的是 Sullivan 方法——引物旁侧在同源区内含子 1 有 6bp 缺失，扩增产物 X－染色体产生 106bp，Y－染色体产生 112bp 片段。因此女性 DNA 有一条 106bp 谱带，男性 DNA 有 106bp 和 112bp 谱带。只用一对引物即可以同时扩增 X、Y 区域，而且拷贝数相同，提供一个快速性别检验方法，灵敏度高，对降解样品分析能力强，特别适合于物证样品性别检验，目前复合扩增体系中常用此引物。

（1）Sullivan 方法

引物 A 5′-CCC TGG GCT CTG TAA AGA ATA GTG-3′

B 5′- ATC AGA GCT TAA ACT GGG AAG CTG-3′

PCR反应 在 50μl 反应体系中，含有 10mmol/L Tris-HCl（pH8.0），50mmol/L KCl，1.5mmol/L $MgCl_2$，175mmol/L dNTPs，0.1μmol/L 每种引物，2μl DNA 提取液，1.5u Taq DNA 聚合酶。

扩增循环参数 94℃ 1min，57℃ 30s，72℃ 1min，35 个循环。

扩增产物检测 PCR 扩增产物在 10% 非变性聚丙烯酰胺凝胶（16 cm×20 cm×0.1 cm）电泳分离，电压为 200V，电泳时间为 1h。电泳后的凝胶在 EB 或 SYBR Green 核酸染色液中染色 30min，254nm 紫外灯下观察结果。或银染染色检测。用荧光染料标记其中一引物，可以与其他 STR 复合扩增，检测。

（2）Haas-Rochholz（1997）设计引物，分析外显子 3 的 3 bp 缺失和内含子 2 的 1 bp 缺失，获得片段更小的 PCR 产物。位于 Amel X 1446-1466 和 Amel Y 1263-1283 引物序列 C 与引物 D（位于 Amel X 1712-1736 和 Amel Y 1533-1557）扩增出 Amel X 80 bp 片段，Amel Y 片段为 83 bp。因此，女性 DNA 有 80 bp 片段，男性 DNA 有 80 bp 和 83 bp 片段。引物 E（引物序列位于 Amel X 1446-1466 和 Amel Y 1263-1283）和引物 F（序列位于 Amel X 1446-1466 和 Amel Y 1263-1283）扩增出 Amel X 的 90 bp 片段，Amel Y 片段为 91bp。女性 DNA 只有 90 bp 片段，男性 DNA 有 90 和 91bp 片段。

①外显子 3 的 3 bp 缺失

引物 C 5′-CCC TTT GAA GTG GTA CCA GAG CA-3′

D 5′-GCA TGC CTA ATA TTT TCA GGG AAT A-3′

PCR扩增 在 50μl 反应体系中，含有 10mmol/L Tris-HCl（pH8.0），50mmol/L KCl，1.5mmol/L $MgCl_2$，200mmol/L dNTPs，0.1μmol/L 每种引物，0.1～10ng DNA，1.25u Taq DNA 聚合酶。

循环参数 95℃ 1min，55℃ 30s，72℃30s，30 个循环。

②内含子 2 的 1 bp 缺失

引物 E 5′-TGG ATG TAA ACA CAG TGC CTG-3′

F 5′-CAT CTT GAG TGT AAG AGT CCG ATA G-3′

PCR扩增 在 50μl 反应体系中，含有 10mmol/L Tris-HCl（pH8.0），50mmol/L KCl，1.5mmol/L $MgCl_2$，200mmol/L dNTPs，0.1μmol/L 每种引物，0.1～10ng DNA，1.25u Taq DNA 聚合酶。

循环参数 95℃ 1min，55℃ 30s，72℃30s，30 个循环。

第十八章　概率统计学基础与 DNA 遗传标记系统参数计算

18.1 概率统计学基础

18.1.1 概率

1. 事件

在综合条件实现时，实验结果中发生的现象或现象可能发生的某一结果，叫做事件（event）。从概率角度，事件可分为以下三种：

必然事件（necessary event）：在一定条件下或在某一试验中，必然会出现的事件称必然事件，记作 U。

不可能事件（impossible event）：在一定条件下不会出现的事件叫不可能事件，用 V 表示。

随机事件（random event）：在一定条件下可能出现也可能不出现的事件称为随机事件，也叫偶然事件，通常用 A 表示。

小概率事件：发生的概率小于 0.05 或 0.01 的事件称为小概率事件。

事件的和：事件 A 和事件 B 至少发生一个所构成的事件称为事件 A 和 B 之和，概率上记作 $A \cup B$，它表示或 A 发生或 B 发生。如“样品为女性或男性”这一事件，即表示为女与男二事件之和。类似的事件 A_1，A_2，A_3，$\cdots A_n$ 至少有一个发生构成的事件称为 A_1，A_2，A_3，$\cdots A_n$ 的和，记作 $A_1 \cup A_2 \cup A_3, \cdots \cup A_n$。

事件的积：事件 A 和事件 B 同时发生所构成的事件称为事件 A 和 B 之积，概率上记作 $A \cap B$，或 AB。类似的事件 A_1，A_2，A_3，$\cdots A_n$ 同时发生构成的事件称为 A_1，A_2，A_3，$\cdots A_n$ 的积，记作 $A_1 \cap A_2 \cap A_3, \cdots \cap A_n$。

不相容事件：事件 A 与事件 B 不能同时发生，即 A 与 B 的积事件是不可能事件，$A \cap B = V$，就称 A 与 B 互不相容或互斥。

独立事件：如事件 B 的发生或不发生，不影响事件 A 发生的可能性，则称事件 A 对于事件 B 是独立的。反之，如事件 B 的发生影响了事件 A 的发生，则称事件 A 对于事件 B 是相依的。

2. 概率（probability）的基本概念与三个定律

（1）概念

设在 n 次试验中，事件 A 出现 m 次，那么把事件 A 出现的次数 m 与试验的次数 n 之比 m/n，称为 n 次试验 A 出现的频率。

在大量重复试验下，设试验的次数为 n，事件 A 出现的次数为 m，如果事件 A 出现的频率 m/n 总是在某个常数 p 附近摆动。就把这个常数 p 称为事件 A 的概率。记为

$$P(A) = p$$

概率是可能性的定量计量，用数值表示某一事件发生的可能性大小。

概率在统计学上用P表示，如$P\leqslant 0.05$，表示某事件出现的概率小于或等于百分之五。

概率可分为先验概率和经验概率两种，先验概率（prior probit）指由定义、理论或规律确定的某事件出现的概率，而经验概率（emperical probit）则是指通过实践活动，由调查、统计资料推断估计出某件事件出现的概率。

小概率原理：如果假设的条件下能够准确地算出事件A出现的概率α为很小，则在假设条件下的n次独立重复试验中，事件A将按预定的概率发生，而在一次试验中则几乎不可能发生。统计学中，常把概率小于0.05或0.01的概率作为小概率。如果计算的概率大于0.05或0.01的概率，则认为不是小概率事件，H_0的假设可能是正确的，应该接受，同时否定H_A；反之，所计算的概率小于0.05或0.01的概率，则否定H_0，应该接受H_A。

（2）概率的基本性质（第一定律）

①任何事件A的概率P（A）总是介于0与1之间：$0\leqslant P(A)\leqslant 1$

②必然事件U的概率等于1，即$P(U)=1$

③不可能事件V的概率等于0，即$P(V)=0$

（3）第二定律（概率的加法原理，addition rule for probability）

如果事件A和事件B互斥：$AB=V$，则和事件$A+B$的概率等于事件A的概率与事件B的概率之和。即

$$P(A+B)=P(A)+P(B)$$

如果出现A，B，C…，共n项事件，其概率为：

$$P(A+B+C\cdots+n)=P(A)+P(B)+P(C)+\cdots+P(n)$$

逆算公式（对立事件的概率）：对于对立的两件事件A与$\overline{A}$，其中一个事件的概率等于1减去另一个事件的概率，即

$$P(A)=1-P(\overline{A})$$

（4）第三定律（概率的乘法定律，multiplication rule for probability）

二个独立事件同时发生的事件的概率为A事件单独发生的概率和B事件单独发生的概率乘积，即

$$P(AB)=P(A)\times P(B)$$

如果同时有几个独立事件，这些独立事件同时发生的概率为它们分别出现的概率乘积，即

如果出现A，B，C…，共n项事件，其概率为：

$$P(A、B、C\cdots n)=P(A)\times P(B)\times P(C)\times\cdots\times P(n)$$

3. 条件概率的概念

设A与B是某试验中的两个事件，在A发生的条件下，B发生的概率称为A发生下B发生的条件概率，记为$P(B|A)$。

对于一个试验中的任意两个事件A与B，$P(A)>0$，$P(B)>0$，事件AB的概率积等于A事件（或B）的概率与A（或B）发生下B（或A）的条件概率之积，即

$$P(AB)=P(A)P(B|A)$$

或 $P(AB)=P(B)P(A|B)$

当有n个事件A_1，A_2，A_3，…A_n，且$P(A_1A_2A_3\cdots A_n)>0$

则

$P(A_1A_2A_3\cdots A_n) = P(A_1)P(A_2|A_1)P(A_3|A_1A_2)\cdots P(A_n|A_1A_2A_3\cdots A_{n-1})$

4. 全概率公式

如果事件 A_1，A_2，A_3，…，A_n 满足：（1）两两相斥，即 $A_iA_j = V$（i，j = 1，2，…，n，i≠j）；（2）和事件是必然事件，即 A_1，A_2，A_3，…A_n = U，则对于任意事件 B 都有下列公式成立：

$P(B) = P(A_1)(B|A_1) + P(A_2)P(B|A_2) + \cdots P(A_n)P(B|A_n)$

上式称为全概率公式。

5. 逆概率公式

如果事件 A_1，A_2，A_3，…，A_n 满足：（1）两两相斥，即 $A_iA_j = V$（i，j = 1，2，…，n，i≠j）；（2）和事件是必然事件，即 A_1，A_2，A_3，…A_n = U，则对于任意概率不为 0 的事件 B 有下列公式成立：

$$P(A_i|B) = \frac{P(A_i)P(B|A_i)}{\sum_{k=1}^{n}P(A_k)P(B|A_k)} \qquad (i = 1, 2, \cdots, n)$$

上式称为逆概率公式（也称贝叶斯公式）。

6. 概率分布

(1) 二项分布

二项分布（binomial distribution）就是二项式展开。二项分布是一种离散型随机变量的分布，例如某个形状，常常可以把其资料分成两个类型，这两个类型的概率分布称为二项分布。非此即彼事件所构成的总体被称为二项总体，其分布称为二项分布。

二项总体的共同特征是：

① 每次试验只有两个对立结果，记作 A 与 $\bar{A}$，它们出现的概率分别为 p 与 q（q = 1 − p）。

② 试验具有重复性和独立性。

以 x 表示在 n 次试验中事件 A 出现的次数。x 是一个离散型随机变量，它的所有可能取值为 0，1，2，…，n，其概率分布函数为：

$P(x) = C_x^n p^x q^{n-x}$

其中，

$$C_x^n = \frac{n!}{x!(n-x)!}$$

P（x）被称为随机变量 x 的二项分布，记作 x（n，p）

二项分布的概率累积函数可用下式表示：

$F(x) = \sum_{x=0}^{n}P(x)$

(2) 泊松分布

在生物学研究中，有许多事件出现的概率很小，而样本容量或试验次数却往往很大。此时，二项分布就变成另外一个特殊的分布，即泊松（Poisson）分布。泊松分布也是一种离散型随机变量的分布，其分布的概率函数为：

$$P(X) = \frac{e^{-\lambda}\lambda^{x}}{x!}$$

式中，λ 为参数，$\lambda = np$，$x = 0, 1, 2, \cdots$。

泊松分布的平均数（μ）、方差（δ^2）、标准差 δ 为：

$\mu = \lambda$

$\delta 2 = \lambda$

$\delta = \sqrt{\lambda}$

（3）正态分布

正态分布也称高斯（Gauss）分布，是一种连续型随机变量的理论分布。它的分布是多数变量都围绕在平均值左右，由平均值到分布的两侧，变量数减少。正态分布是一种在统计理论和应用上最重要的分布。实验误差的分布一般服从于这种分布，许多生物现象的计量资料均近似服从这种分布。同时，在一定条件下，正态分布还可以作为离散型随机变量或其连续型随机变量的近似分布。当 n 相当大或 p 与 q 基本接近时，二项分布接近于正态分布，当 λ 较大时，泊松分布也接近正态分布。

正态分布的概率函数可由二项分布的概率函数在 $n \rightarrow \infty$ 时导出，其方程为：

$$f(x) = \frac{1}{\delta\sqrt{2\pi}} e^{-1/2\left[(x-\mu)^{1/\delta}\right]^2}$$

式中，f（x）为正态分布的概率密度函数，表示某一 x 值出现的概率密度函数值，μ 为总体平均数，δ 为方差。.

18.1.2 **统计基础**

1. 常用统计学术语

总体与样本　具有相同性质的个体所组成的集合称为总体。按总体单位的数目，总体可分为有限总体和无限总体。个体极多或无限多的总体称为无限总体，个体有限的总体称为有限总体。

样本是指从总体中抽取的一部分个体，是总体的一部分，它是总体的代表。在一个样本内可以包含有不同的个体数，样本数包含的个体数目称为样本含量。根据样本大小，可将样本分为大样本和小样本两种。大样本是指含量在 30 个个体以上，30 个个体以下的样本称为小样本。

参数与统计数　参数也称参量，是对一个总体特征的度量。如总体平均数、总体标准差等均为参数。因为总体一般都很大，有的甚至不可能取得，所以总体参数一般不可能计算出来。可以通过对总体抽取样本，统计样本的统计数来估计总体参数，从样本中计算所得的数值称为统计数，它是总体参数的估计值。

变量与常数　相同性质的事物间表现差异性或差异特征的数据称为变量或变数，它是表示在一个界限内变动着的性状的数值。如某一基因座在人群中有等位基因 6，7，8，9，10，11，12，13 等 8 个变量，记作 x_i（i = 1，2...，8），表示 x_1 到 x_8 之间任一数值；亦称 x_i 为随机变量。随机变量（random variable）是一种由测验结果而定，并且随着试验结果而变的变量。

变量按其性质可分为连续变量和非连续变量。连续变量表示在变量范围内可抽出某一

范围的所有值，这种变量之间是连续的、无限的，如 RFLP 图谱中的 DNA 片段大小为连续变量。非连续变量，也称为离散变量，表示在变量数列中，仅能取得固定数值。

变量可以是定性的，也可以是定量的。定性的变量往往表示某个体属于几种互不相容的类型中的一种，如人的血型有 A、B、AB 和 O 型，等等。定量的变量是指可测量的，如电泳图谱中的带数等。

常数 表示能代表事物特征和性质的数值，通常由变量计算而来，并且在一定的过程中是不变的，如某样本平均数标准差、变异系数等。

零值假设（null hypothesis） 零值假设也称无效假设，用 H_0 表示。在对两个样本统计进行差异显著性测定时，统计学都要提出一个假设，总是把两个样本间表现出的差异认为是来自同一个总体。样本间没有根本性差别。假设差异是由抽样误差所致，然后进行统计分析，依据实验所得数据，按照一定概率标准来推断这种假设是否成立，从而判断两个样本统计差异的真假，得出统计最显著性测定的结论。

根据研究问题的对象不同，假设可以用不同形式提出。如测定二个人群的差异性，可先假设他们之间无差异，然后按实际数据计算这种假设出现的概率。依概率的大小来判断假设的真假，确定二事物间差异的真假。

置信限度（confidence limit） 置信限度也称可信限度（区），是指在一定概率下，用样本均数估计总体均数可能的范围，即总体均数的可能区间范围。由于样本均数是一个变量，不能确切确定总体均数。因此，比较合理的是给总体参数一个范围，即在一定概率下，总体均数可能在多么大一个范围内。

自由度（degree of freedom） 是指能够自由活动的变量的个数，也指可以自由活动的范围。从几何学上的点来看自由度，在平面上的一点，通常用（x、y）表示，如果没有条件限制，这个点可以沿着纵横二轴任意移动，其自由活动的范围有两个。如果有一个条件限制它，如 $x+y=5$ 时，x 和 y 就只有一个能任意选择，即当 $x=2$ 时，y 只能等于 3，且（x，y）点只能在 $x+y=5$ 这条直线上活动。故自由度只有一个。在空间的一个点是由三个数字表示的，即（x，y，z），如不加以限制，它有三个自由度，可以向三个方向变动，如果使 $x+y+z=10$，这个空间点只有两个数可以自由活动，自由度就只有两个。因此，每加上一个条件，自由度就减少一个。由此类推，数学上把 x_1、x_2、$x_3 \cdots x_n$ 作为几度空间里的点，如无条件限制，这个点就有几个自由度，如加上下面一个条件，即

$$x_1+x_2+x_3+\cdots+x_n=\sum x$$

有了这个条件的限制，自由度就减少一个，成为 $n-1$。如计算样本标准差时因受样本估计常规数均数的限制，其自由度就为变数的个数减去 1，即 $n-1$。如果同时存在两个或三个条件的限制，则自由度 n 相应的减少 2 或 3、即为 $n-2$，$n-3$。如计算相关、回归有关统计量时，自由度受到两个估计常规数的限制，其自由度为 $n-2$。

生物统计学之所以应用自由度，是为了在小样本情况下，不降低其精确程度，因为通常计算采用的小样本只是总体中极小的一部分，其全距远比总体的全距为小。这样用样本估计总体，对总变量平均时，不用样本包含个数 n，而用自由度 $n-1$ 去除。

错误与机误（error） 错误指工作人员在试验中疏忽大意或有意识的加大、减少数据等造成的差错，使试验结果不真实，如在试验中称量药物时未校正天平，这是人为的差

错，可以避免。

机误是指试验中出现的差异是由机会造成的。如在随机抽样中，出现较大或较小的数据较多，这是由于总体中个体间存在着差异，由于机会作用，出现差异，它是不可避免的，只能设法减小、而不能完全消灭。增加次数或加大含量可以减小，次数越多，误差越小。

百分比与百分率（percentage）　百分比与百分率为两种不同统计指标。百分比是结构指标，表示事件占的比例，即部分对全体之比；百分率则是频率指标，表示事件出现的频率。

2. 方差与标准差

为了度量其变异程度，对含有 n 个观测数 x_1，x_2，$\cdots x_n$ 的样本，可以用各观测数离均差的大小来表示，但由于 $\sum(x-\bar{x})=0$，不能反映其变异程度。若将离均差先平方再求和，即 $\sum(x-\bar{x})^2$，就可以消除上述弊病。但这样还有一个缺点，就是离均差平方和常随样本容量大小而改变。为了便于比较，用样本容量 n 来除离均差平方和，得到平均的平方和，简称方差或均方。对于样本来说，其方差 s^2 为：

$$s^2=\frac{\sum(x-\bar{x})^2}{n-1}$$

对于总体，其方差 σ^2 为：

$$\sigma^2=\frac{\sum(x-\mu)^2}{N}$$

式中：n－1 为自由度，N 为有限总体容量

方差的平方根值就是标准差。

样本的标准差为：

$$s=\sqrt{\frac{\sum(x-\bar{x})^2}{n-1}}$$

总体标准差 σ 为：

$$\sigma=\sqrt{\frac{\sum(x-\mu)^2}{N}}$$

3. 统计推断——假设检验的原理与方法

假设检验就是根据总体的理论分布和小概率原理，对未知或不完全知道的总体提出两种彼此对立的假设，然后由样本的实际结果，经过一定的计算，作出在一定概率意义上应该接受的那种假设的推断。生物统计学，一般认为小于 0.05 或 0.01 的概率为小概率。通过假设检验，可以正确分析处理效应和随机误差，作出可靠的结论。

假设检验的步骤：

（1）假设　检验假设首先对总体提出二个假设，一个是无效假设记作 H_0，另一个是备择假设，记作 H_A。无效假设是直接检验的假设，是对总体提出的一个假想目标。

（2）确定显著水平　进行无效假设和备择假设后，要确定一个否定 H_0 的概率标准，这个概率标准叫显著水平，记作为 α。α 是人为规定的小概率界限，一般 $\alpha=0.05$ 和 $\alpha=0.01$ 两个显著水平。

（3）概率　在正确的前提下，根据抽样分布的统计数，进行假设检验的概率统计。

（4）推断是否接受假设　根据小概率原理作出是否接受 H_0。

对离散型资料的假设检验，通常都采用 χ^2 检验。一般有两种类型。一类是适合性检验，这种方法对样本的理论数先通过一定的理论分布推衍出来。然后用实际观测值与理论数比较，从而得出实际观测值与理论数之间是否吻合。因此适合性检验也叫吻合度检验。另一类是独立性检验，是研究两个或两个以上属性的计数资料或属性资料间是相互独立的或是相互联系的，这时可以假设所观测的各属性之间没有关联，然后证明这种无关联的假设是否成立。一个样本的同质性检验，也需进行 χ^2。

4. χ^2 检验

χ^2 的原意是从方差为 σ^2 的正态总体中，随机抽取含量为 n 的样本，算出样本方差 s^2。在研究样本方差的分布时，通常将它标准化，得到一个不带有任何单位的纯数。

$$\chi^2=\sum_{i=1}^{k}\left[(x-\mu)/\sigma\right]^2=\frac{1}{\sigma}\sum_{i=1}^{k}(x-\mu)^2$$

对计数资料和属性资料进行 χ^2 检验，其基本原理是应用理论推算值与实际观测值之间的偏离程度来决定其 χ^2 值的大小。理论值与实际值之间偏差越大，越不符合；偏差越小，越趋于符合，若两值完全相等时，表明理论值与实际值完全符合。在计算理论推算值 E 与实际观测值 O 之间的符合程度时，最简单的方法是比较两者差数的大小。但由于 O－E 有正有负，则 $\sum(O-E)$ 趋近于零，不能真实地反映理论推算值与实际观测值差值的大小，故采用 $\sum(O-E)^2$，这样就可以消除负号。实际观测值与理论推算值相差越大，则也 $\sum(O-E)^2$ 越大，反之亦然。由 $\sum(O-E)^2$ 似乎可以度量观测值与理论值的相差程度，实际上这个绝对差异数还不足以表示相差程度。因此，采用

$$\sum\frac{(O-E)^2}{E}$$

使其转化为相对值，这个值便是 χ^2 值，即：

$$\chi^2=\sum_{i=1}^{k}\frac{(O_i-E_i)^2}{E_i}$$

式中：O 为实际观测值，E 为理论推算值。

这是随自由度 df＝k—1 而变化的连续型分布。

由公式可知 χ^2 最小值为 0，随着 χ^2 值的增大，观测值与理论值符合度越来越小。所以 χ^2 的分布是由 0 到无限大的变数。实际上其符合程度由 χ^2 概率决定。由 χ^2 值表可知，χ^2 值与概率 P 成反比。χ^2 值越小，P 值越大；χ^2 越大，P 值越小。因此，由 χ^2 分布对计数资料或属性资料进行假设检验。χ^2 检验的步骤为：

（1）提出无效假设 H_0：观测值与理论值的差异由抽样误差引起，即观测值＝理论值。同时给出相应的备择假设 H_A：观测值与理论值的差值不等于 0，即观测值≠理论值；

（2）确定显著水平 α，一般可确定为 0.05 或 0.01；

（3）计算样本的 χ^2：求得各个理论次数 E，并根据各实际次数 O_i，代入上式中，计算样本的 χ^2。

（4）进行统计推断。求得 df＝k－1，从 χ^2 表中查出 χ^2_α 值。如果实得 $\chi^2<\chi^2_\alpha$，即表明

$P>\alpha$，应接受 H_0、否定 H_A，表明在 α 显著标准下理论值与实际值差异不显著，二者之间的差异系由抽样误差引起。如果实得 $\chi^2>\chi^2_\alpha$，即表明 $P<\alpha$，应否定 H_0，接受 H_A，则表明在 α 显著标准下理论值与实际值差异是显著的，二者之间的差异是真实存在的。

由于 χ^2 分布是连续的，而计数资料是离散的，故所得的 χ^2 值是一个近似值。为了使离散型的计算结果适合于连续型分布给出的概率，在计算 χ^2 时应注意以下两个问题：

(1) 任何一组的理论次数 E_i 都必须大于5，如果 $E_i\leqslant5$，则需要并组或增大样本容量，以满足 $E_i>5$；

(2) 在自由度 $df=1$ 时，需进行连续性矫正，其矫正的 χ^2_α 为：

$$\chi^2_C=\sum_{i=1}^{k}\frac{(O_i-E_i)^2}{E_i}$$

对同一资料，进行矫正的 χ^2 值要比未矫正的 χ^2 值小。当自由度 $d\geqslant2$ 时，由于 χ^2 与 χ^2_C 相差不大，所以一般不再进行连续性矫正。

应用 χ^2 测验，要强调的是实验求出的 χ^2 值在一定的自由度下大于 χ^2 表相应 $\chi^2_{0.05}$ 值时（即 $P<0.05$），被认为是差异显著。这个差异显著只是指实验所得的数据与理论数据有显著差异，其"显著"差异是指概率来说的，与我们日常所说的显著在含义上是不完全相同的。同样，差异不显著也是指概率而言的，并非事物没有差异。通常出现差异不显著，可能有两种情况，一种是实际数据和理论数据不属于同一总体，且二总体是有差异的，由于样本采样有限，其内部规律性的差异被偶然性的变异所掩盖，未能表现出来，如果是这种情况，就得进行重复实验，严格控制影响实验的条件，采取精密的实验仪器，精心对实验进行设计或增加观察的例数，就可以得出合理的结论。另一种情况是实际次数与理论数据的总体分布是一致的，在这种情况下也不会得出差异显著的结论。根据实验判断，一般P值大于或接近0.05时，重复实验容易得出显著的结论；当P值界于0.20到0.05时，实验需要重复；如 $P>0.20$ 时，这种把握性就比较大，没有必要进行重复实验，可以断然做出差异不显著的结论。

其次，应用 χ^2 方法检验实验结果时，首先要检验每格内的预期数值（理论次数）是否够大，一般不应少于5，否则，容易导致错误的结论。对于用百分数表示的资料，应把百分数按总数换算成数值频数，即采用绝对数值，然后计算差数。而在四格表计算中如有理论值小于5或者实际数出现零时，用上面的公式求值是有偏倚的，特别是总观察值小于40时，则不能用上公式计算，而要用精确检验法，直接计算概率以作判断，其计算方法为

$$P=\frac{(a+b)!\ (c+d)!\ (a+c)!\ (b+d)!}{a!\ b!\ c!\ d!\ n!}$$

5. 适合性检验

比较观测数与理论数是否符合的假设检验叫适合性检验，也称吻合性检验。例如，在遗传学上，常用 χ^2 检验来测定所得的结果是否符合哈-德温格平衡等。许多与已有理论比率进行比较的资料，也需用 χ^2 来作适合性检验。适合性检验是 χ^2 检验最常用的方法。

作适合性检验时，可提出无效假设 H_0：$O-E=0$，即认为观测数与理论数之间没有差异，再计算样本 χ^2 值，根据规定的显著性水平 α 和自由度 df 从 χ^2 值表中查出 χ^2_α，当

$\chi^2 > \chi_\alpha^2$时，拒绝 H_0；接受 H_A；$\chi^2 < \chi_\alpha^2$ 时，接受 H_0。

6. 独立性检验

独立性检验是研究两个或两个以上因子彼此之间是独立的还是相互影响的类统计方法。例如，同一染色体上的两个基因座有无关系，若无关系则说明两者是独立的。若有关系，说明两个基因座是连琐的。这种关系是否达到显著呢？可以进行 χ^2 检验。具体做法：提出无效假设，根据无效假设计算理论数，在一定的自由度下以给定的显著性水平做出推断。若拒绝无效假设，则说明两种因子之间的关联是显著的；若接受无效假设，则说明两种因子之间无关联，是相互独立的。独立性检验的形式有多种，常利用列联表进行检验。

(1) 2×2 列联表的独立性检验

设 A，B 是一个随机试验中的两个事件，其中 A 可能出现 r_1、r_2 个结果，B 可能出现 c_1、c_2 个结果，两因子相互作用形成 4 格数，分别用 0_{11}，0_{12}，0_{21}，0_{22}表示，表 18－1 即为 2×2 列联表的一般形式。

表 18－1　2×2 列联表的一般形式

行＼列	c_1	c_2	总和
r_1	0_{11}	0_{12}	$R_1 = 0_{11} + 0_{12}$
r_2	0_{21}	0_{22}	$R_2 = 0_{21} + 0_{22}$
总和	$C_1 = 0_{11} + 0_{21}$	$C_2 = 0_{12} + 0_{22}$	T

2×2 列联表的 χ^2 检验需经以下步骤：

①提出无效假设 H_0：事件 A 和事件 B 无关，即事件 A 和事件 B 相互独立、同时给出 H_A：事件 A 和事件 B 有关联；

②给出显著水平 α；

③依据自由度，可以推算出理论数，计算 χ^2 值；$\chi^2 = \frac{(|0_{11}0_{22} - 0_{12}0_{21}| - T/2)^2\, T}{R_1R_2C_1C_2}$

④确定自由度，df＝（r－1）（c — 1），或者写为（行— 1）（列— 1），进行推断。若所计算的 $\chi^2 > X_\alpha^2$，则 P ＜ α，表明实际观测数与理论数不一致，应否定 H_0，接受 H_A；若 $\chi^2 < \chi_\alpha^2$，则 P＞α，表明实际观测数与理论数相一致，应接受 H_0，否定 H_A。

(2) 2×c 列联表的独立性检验

2×c 列联表，即横行分为两组，纵行分为 c≥3 组的资料，2×c 列联表的一般形式如表 18－2 所示。

表 18－2　2×c 列联表的一般形式

行列	1	2	c	合计
1	0_{11}	0_{12}	0_{1c}	R_1
2	0_{21}	0_{22}	0_{2c}	R_2
合计	C_1	C_2	C_c	T

2×c 列联表理论值的计算和 2×2 列联表一样，自由度 df＝（2－1）（c－1）＝c－1，

由于 $c \geqslant 3$，故 $df \geqslant 2$，因此计算 χ^2 时不需要做连续性矫正。

$$X^2 = \frac{T^2}{R_1 R_2}\left[\sum \frac{O_{1i}^2}{C_i} - \frac{R_1^2}{T}\right]$$

(3) r×c 列联表的独立性检验

r×c 列联表是指 r≥3、c≥3 的计算资料，表 18－3 是 r×c 列联表的一般形式。

表 18－3　r×c 列联表的一般形式

行　列	1	2	c	合计
1	O_{11}	O_{12}	O_{1c}	R_1
2	O_{21}	O_{22}	O_{2c}	R_2
r	O_{r1}	O_{r2}	O_{rc}	R_r
合计	C_1	C_2	C_c	T

r×c 列联表各项理论频数的计算同 2×2 列联表及 2×c 列联表，可由下式计算：

$\chi^2 = T\ [\sum\ (O_{ij}^2/R_iC_j)\ -1]$

式中：$i=1, 2, \cdots, r$，$j=1, 2, \cdots, c$。自由度 $df=(r-1)(c-1)$

18.2 遗传标记系统多态性的评估

DNA 多态性遗传标记成千上万，每个标记的多态性、法医学应用价值及特点都有可能不同，因此，在应用任何一个遗传标记前，必须对其进行评估。评估包括相互关联的两个方面：一是对于标记系统本身多态性的遗传评估，即通过群体调查，获得各项遗传多态性参数。该项评估的目的是为应用时选择一个好的遗传标记系统，以便获得较高的个人识别和亲权鉴定的效率；二是对于所建立的方法进行法医学应用的有效性的评估。

评价遗传标记多态性的常用参数有：基因频率、基因型频率、杂合度（H）和多态性信息含量（PIC）、识别能力（D）、非父排除率（PE）和平均亲权指数（PI）等。通过群体调查可以获得标记的参数。调查个体量必须足够大，如 STR 基因座必须在 200 个体以上。计算群体遗传学参数，证实调查结果符合 Hardy－Weinberg 平衡，并反映一个标记系统在特定人群中的多态性。

18.2.1 等位基因频率和基因型频率

等位基因频率（allele frequency）是指在一个二倍体的某特定基因座上某一个等位基因占该基因座上等位基因总数的比率，这是群体遗传结构的一个最基本的尺度，该基因座上所有等位基因的频率之和为 1。对于法医学应用的大多数 DNA 遗传标记，前面已提到它不是基因和等位基因，而是等位片段，只是借用了基因和等位基因它们的概念。

基因型频率（genotype frequency）指的是群体中某特定基因型个体的数目占个体总数的比率，该基因座上所有基因型频率之和为 1。等位基因频率、基因型频率是群体的遗传结构表达方式。

群体遗传用群体中各等位基因的比率或各种基因型的比率来表示群体的遗传组成，对一个群体而言，某个基因座出现两种或两种以上的等位基因，每个等位基因频率都达到

0.02 以上，则称这个群体在这个基因座上具有遗传多态性（genetic polymorphism）。等位基因频率 0.02 是区别多态和突变的人为标准。也有文献报道以 1%作为标准的，即认为某个基因座出现两种或两种以上的等位基因，基因频率超过 1%时该群体在该基因座具有多态性。

1. 等位基因频率计算

对于离散型等位基因频率常用计数法计算。比如 STR 基因座：

等位基因 i 的频率 p 为：

$$p_i = \frac{n_i}{2N} = \frac{2n_1 + n_2}{2N}$$

式中：

n_1——含等位基因 i 的纯合子个体数；n_2——含有该等位基因杂合子个体数；

N——为调查的群体个体数。

该基因座所有等位基因频率的总和为 1，即 $\Sigma p_i = 1$。

对于非连续等位基因频率分布的频率计算，应用较为广泛的有两种方法，即固定箱和滑动箱法计算。固定箱法采用“装箱”合并的方法，将一系列等位基因片段按从大到小的顺序分成若干组，以不连续的分组方式将连续值处理成不连续的数据群，将连续变量基因频率处理成不连续基因频率分布。其原则是将相邻的多态片段合并为一组，当作一个等位基因看待，计算合并后的组的频率。经装箱处理的单基因座探针识别的基因座等位基因频率又称“箱频率”或“固定箱频率”。

2. 基因型频率

基因型频率也可以由记数法获得，总的基因型频率为 1。基因型的理论值可以根据 Hardy - Weinberg（HW）平衡推导的结论计算，纯合子频率为等位基因频率的平方（p^2），杂合子频率为两等位基因频率的乘积的 2 倍（2pq）。计数法是一种准确的方法，不涉及估计问题。

纯合子基因型的预期值 =（基因频率）^{2} × 样本含量 = $p^2 \times N$

杂合子基因型的预期值 = 2 ×（基因 1 频率）×（基因 2 频率）× 样本含量 = $2pq \times N$

18.2.2 杂合度

一个遗传标记的多态性程度通常用杂合度定量评估。杂合度（heterozygosity，Het）指群体中杂合子的频率，常用 H 来表示杂合度。它是估计群体内遗传变异的一种简单方法。实际杂合度或观察杂合度也可以由记数法直接获得：

$$H_{观察值} = 杂合子个数/群体个体总数$$

理论杂合度或期望值根据 Hardy - Weinberg 平衡可用下式计算：

$$H_{理论} = 1 - \sum_{i=1}^{n} p_i^2$$

式中：

p_i 代表群体中第 i 个等位基因的频率，在 Hardy - Weinberg 平衡状态下，群体中纯合子的频率为 Σp_i^2，而杂合子频率为 1 减纯合子的频率。

等位基因数越多，p_i 越接近平均数（1/n），杂合度就越大。

18.2.3 多态性信息含量

多态性信息含量（polymorphism information contents, PIC）是评价一个多态性基因座的使用价值的一个量化概念，PIC 是指在某一家系中可以利用基因座多态性作为遗传标记进行基因连锁分析的概率。PIC 是由该基因座的等位基因数、等位基因频率分布两个因素决定。与杂合度一样，标记基因座的等位基因数越多、各等位片段的频率越均匀，PIC 越大。对于常染色体上的多态性基因座的 PIC 的计算公式如下：

设遗传标记等位基因分别为 A1，A2，A3...An，共 n 个等位基因，其等位基因频率分别为 p_1，p_2，$p_3 \cdots p_n$。在 Hardy - Weinberg 平衡状态下，群体中纯合子的频率为 $p_1^2 + p_2^2 + p_3^2 \cdots + p_n^2$，杂合子 A_iA_j 频率为 $2p_ip_j$，

$$PIC = 1 - \sum_{i=1}^{n} p_i^2 - \sum_{i=1}^{n} \sum_{j=i+1}^{n} 2p_i^2 p_j^2$$

式中：

p_i 是第 i 个等位基因的频率，第二项排除了不能提供信息的纯合子，而第三项排除了杂合子中半数不能提供信息的杂合子。

常染色体标记的 PIC 小于该基因座的杂合度，X 染色体多态性标记的 PIC 等于杂合度。

18.2.4 匹配概率 Pm

匹配概率 Pm（probability of matching, Pm）或叫偶合机率（probability of coincidence, Pc）指的是人群中随机抽取二无关个体在特定基因座二者的基因型纯粹由于机会一致的机率，匹配概率为各基因型频率的平方和：

$$Pm = \sum_{i=1}^{n} P_i^2$$

式中：

n 为该基因座的基因型数；P_i 为该基因座第 i 个基因型的频率。

若检测 k 个独立的遗传标记，其联合 P_M 值为个各个 Pm 的乘积。

计算公式为：

$$P_M = P_{m1} \times P_{m2} \times P_{m3} \times \ldots \times P_{mk} = \prod_{j=1}^{k} Pj$$

式中：k 为检验的 DNA 遗传标记数

P_{mj} 为检验的第 j 个标记的 Pm。

18.2.5 个人识别力

个人识别力（discrimination power, DP），是指从同一群体中随机抽取两名个体的遗传标记基因型不相同的概率，即遗传标记识别无关个体的能力，也是遗传标记对于个人识别效能的定量评估。对遗传标记而言，从同一群体中随机抽取两名个体表型不相同的概率越高，其识别无关个体的能力就越强。计算 DP 值的公式（Fisher 1951）为：

$$DP = 1 - Pm = 1 - \sum_{i=1}^{n} P_i^2$$

式中：n 为一个遗传标记的基因型数目，P_i 为群体中第 i 个基因型的频率。

若检测 k 个独立的遗传标记，其累积个人识别能力计算公式为：

$$TDP = 1 - (1 - DP_1)(1 - DP_2)(1 - DP_3) \cdots (1 - DP_k) = 1 - P_M$$

$$TDP = 1 - P_1 \times P_2 \times P_3 \times P_4 \times \cdots \times P_k = 1 - \prod_{j=1}^{k} Pj$$

式中 Pj 为第 j 个遗传标记的 Pm 值，ΠPj 为 k 个遗传系统的总 P_M 值。检查数种遗传标记，先按公式求出每种遗传标记的 Pm 值，然后再求出累计 PM 值，最后求出累计 DP 值。表 18－4 列出了汉族群体的 9 个 STR 基因座的个人识别能力。所用遗传标记数目越多，鉴别机率愈高，鉴别能力愈强。

表 18－4 汉族群体 9 个 STR 基因座的个人识别能力及累积个人识别能力

基因座	个人识别能力	累积个人识别能力
D3S1358	0.869	0.869
VWA	0.935	0.9915
FGA	0.964	0.9997
D8S1179	0.956	0.9999866
D21S11	0.939	0.9999992
D18S51	0.960	0.999999968
D5S818	0.901	0.9999999969
D7S820	0.930	0.99999999979
D13S317	0.931	0.999999999987
TPOX	0.789	>0.99999999999
TH01	0.833	>0.99999999999
CSF1PO	0.858	>0.99999999999
D16S539	0.921	>0.99999999999

对于连锁的遗传标记，由于它们紧密连锁，它们的累积个体识别率不能按上述独立基因座方法连乘计算，如 Y－染色体上的基因座、线粒体的碱基序列差异，应统计出几个基因座或区域的单倍型后，计算单倍型的频率，然后再计算个体识别率或叫基因差异性（gene diversity，GD）或单倍型的差异性（haplotype diversity）。其计算公式：

$$GD = [n(1 - \Sigma p_i^2)] / (n - 1)$$

式中：p_i 为单倍型频率

对于 X 染色体上的基因座，计算公式：

$$Dp = 1 - 2(\sum_{j=1}^{k} p_j^2)^2 + \sum_{j=1}^{k} p_j^4$$

18.2.6 非父排除概率

非父排除概率，是指不是孩子生父的男子能被遗传标记排除的概率。它是遗传标记对于亲权鉴定效能的定量评估。不是孩子生父的男人而被控为父亲时，理论上可以根据遗传标记否定其为生父。但在遗传标记的鉴别能力较差时，没有血缘关系的男人与孩子的遗传标记有时也会相同，偶然符合孟德尔遗传规律，从而不能否定该男子与孩子的亲缘关系。因此对于法医学应用的遗传标记必须知道父权排除概率（Excluding probability of paternity，EP），确切地说非父排除概率是：对于不是孩子生父而被控为生父的男子，用遗传标记否

定其父权有多大的可能性。

设 p_i 代表群体中第 i 个等位基因频率，p_j 代表群体中第 j 个等位基因频率，并且等位基因 i 不等于等位基因 j，则排除概率为：

$$EP = \sum_{i=1}^{n} p_i \ (1 - pi)^2 - \frac{1}{2} \ (\sum_{i}^{n-1} \sum_{j \neq i}^{n} p_i^2 p_j^2 \ (4 - 3p_i - 3p_j))$$

式中：n 为等位基因数。

K 个独立的 DNA 遗传标记系统的累积非父排除率（cumulative exclusion power，CEP），计算公式：

$$CEP = 1 - \ (1 - EP_1)\ (1 - EP_2)\ (1 - EP_3)\ \ldots\ (1 - EP_k) \ = 1 - \prod_{i=1}^{k} \ (1 - EP_i)$$

式中：EP_k 为第 k 个 DNA 遗传标记的 EP 值。检查多种遗传标记，按各种遗传标记的遗传方式求出 EP 值后，再按公式求出总的 CEP 值。所用遗传标记数目越多，累积非父排除概率愈高，鉴别能力愈强。

18.2.7 Hardy - Weinberg **平衡的验证**

Hardy - Weinberg 定律，是指人群中的基因型和等位基因频率按一定比率分布，并且在向子代传递中保持稳定，世代不变。符合该平衡定律的实际人群应满足以下要求：随机婚配，无突变或人群中的基因突变率保持恒定；无大规模迁移导致人群结构改变现象；无针对某一基因型的选择；人群数足够大，不因某一偶发事件导致的等位基因频率波动在人群间扩散。只有从处于 Hardy - Weinberg 平衡的群体中获得的群体数据，才能保证其可靠性和准确性，才能应用于法医学个体识别和亲权鉴定。因此，对于新的遗传标记所获得的数据，有必要进行 Hardy - Weinberg 平衡验证，确定此群体是否处于 Hardy - Weinberg 平衡中。常用基因型的预期值和观察值之间是否吻合来检测人群的 Hardy - Weinberg 平衡。

基因型的预期值和观察值之间的吻合程度一般用 χ^2 检验来量度，把每一种基因型的 χ^2 值相加获得总的 χ^2 值，然后查表求出 P 值，一般以 $P > 0.05$ 作为无显著性差异的界限。计算公式为：

$\chi^2 = \sum$（观察值 - 预期值$)^2$/ 预期值

其中 χ^2 检验的自由度为：df = 基因型数 - 等位基因数

此处的基因型数，是指实际观察到的基因型类型数目，而不是根据等位基因数推算出的预期基因型数目。

群体调查的样本数相对过少，而等位基因数目相对过多时，基因座实际观察到的基因型数目大大少于根据等位基因数推算出的预期基因型数目，用基于大样本理论的 χ^2 检验验证 Hardy - Weinberg 不平衡，可以将频率低于 5% 的合并成一组计算。

18.2.8 DNA **标记独立性检测**

DNA 标记独立性，是指不同遗传标记相对独立，遗传过程中不同标记或基因座上的基因可随机组合，标记或基因座间没有统计学意义的关联与连锁。基因座独立性用 2×2 统计表及 χ^2 检验确定。例如：

经 χ^2 检验，两标记系统间 $\chi^2 = 3.41$，$P > 0.05$，表明两标记无明显关联。说明没有统计学意义的关联与连锁，两标记相互独立。用法医物证学技术检测若干遗传标记后，可用概率乘法定律，使样本由一系列表型组成一种稀有的组合表型，可用概率乘法定律计算系

统的各项联合总指标，如偶合机率、识别率等。如果两个标记间存在连锁关系，则不能用乘法计算。一般不在同一条染色体上的标记相互是独立的，位于同一染色体上物理距离大于 10Mb 以上的标记也相互独立，没有连锁关系；位于同一染色体上物理距离小于 10Mb 的标记存在连锁关系，他们不相互独立。

18.2.9 单倍型频率

两基因座间有关联或连锁时，单个基因座不具有独立性，不能用概率乘法定律来计算联合识别率等指标。此时需用单倍型频率来描述群体的遗传结构或计算法医学相关的概率。单倍型指连锁基因座等位基因的组合，构成单倍型的等位基因呈连锁遗传。常染色体遗传标记的单倍型需家系调查才能确定，因为连锁的两个基因座均为杂合子时，不能根据个体分型的结果判断哪两个基因连锁在一起遗传。母系遗传的 mtDNA 和父系遗传的 Y 染色体遗传标记则可根据个体分型的结果判断单倍型。mtDNA 和 Y 染色体遗传标记的单倍型频率可用计数法直接计算出，计算公式为：

$$p_X = \frac{n}{N}$$

式中：

P_x 代表单倍型 X 的频率

n 为单倍型 X 的计数

N 为个体数

18.2.10 多基因座 DNA 指纹图匹配率的计算

由于多基因座 DNA 指纹图中的一条片段，可能来自某纯合子基因座上相同长度的基因，也可能是来自 2 个或 2 个以上基因座的相同长度基因，无法判定某片段来自哪一个多态基因座；无法确定各基因座的基因型，也不能通过群体调查确定等位基因及基因型频率。那么它的匹配概率计算不能用前面介绍的方法计算，可以通过群体调查研究人群中每个片段的平均频率（q）和无关个体共有一个片段的平均概率（x）来计算匹配率。x 值是目前计算指纹图偶合概率的重要基础数据。由平均概率计算出两个比对样品之间的相关指数或偶合率确定他们的关系。详细请参见 DNA 指纹图章节。

18.3 应用评估

18.3.1 家系调查和突变率

为了评价遗传标记能否应用于亲子鉴定中，避免因遗传标记的突变导致错误地排除亲子关系，必须进行家系调查，观察 500 次以上减数分裂，计算基因座的突变率，证实基因座等位基因传递遵循孟德尔遗传定律。

基因突变（gene mutation），是指基因的核苷酸顺序或数目发生改变。突变率，是指每代细胞发生突变的百分率。不同基因座的突变率有差异，大多数基因座突变率很低，个别基因座突变率较高。

突变率的计算公式：

$$\mu = \frac{x}{N}$$

式中：

x——实际观察到的突变个数

N——调查的独立的减数分裂数

国际DNA委员会要求，用于亲子鉴定的遗传标记突变率应在0.2%以下。过高突变率的基因座不宜用于亲子鉴定。

18.3.2 组织同一性测定

对新鲜或一定条件下保存的组织和体液进行检验来评估标记。同一个体的不同组织、器官、分泌物、排泄物的DNA基因型应一致，无器官特异性，这是法医DNA分析可以根据不同组织进行个体同一认定的理论基础。个体组织的同一性不受营养、年龄及环境条件的影响，而保持终生不变。同卵双生子的基因组DNA相同，他们的DNA分型结果也完全一致。

18.3.3 方法应用有效性的验证

1. 种属特异性试验

法医物证检材多是斑痕类，可能含有非人的生物物质如动物、植物或细菌物质，凭肉眼无法判断。如果一个遗传标记不是人特有的，动物DNA也能得到分型，如家禽类也能得到与人一样的DNA指纹图，如不预先进行种属试验，直接进行DNA分型，误将动物血判为人血，有可能错误地排除罪犯，或将与案件无关的人划为嫌疑对象，将案件的侦察与审判工作引入歧途。因此，阐明遗传标记的种属特异性有助于针对性地设计物证检材的分析程序，有利于对结果作出正确解释。

常用DNA遗传标记中，多基因座探针无种属特异性。已发现多种探针可与动物DNA出现交叉反应，例如探针33.6和33.15对狗、猫、鸟类甚至鱼类DNA有交叉反应。3′-HVR探针与牛、马、狗、鹿DNA有交叉反应，虽然杂交带密度较低，但仍然严重干扰鉴定结果。MZ1.3探针本来就是用噬菌体DNA筛选出来的，与常见野生和家养动物DNA均有相当高的同源性，在个人识别时，动物DNA污染的干扰更明显。

单基因座探针大多经过严格的筛选，不仅具有基因座特异性，而且也具有种属特异性。例如，Tajima等（1989）研究了2.3 kb串联重复序列探针（pH 12）的种属特异性，该探针能与人DNA杂交，与蛙、鱼、老鼠、老鼠、兔子、鸡、猫、狗、羊以及细菌DNA不能杂交，但与日本猴和黑猩猩DNA有杂交信号。

各STR基因座引物设计主要根据人基因组DNA序列设定，一般对人的特异性都较好，但有些基因座并非人类特有。某些生物尤其是灵长类动物也可能有扩增产物，因此对一个新基因座体系需进行种属特异性研究。Crouse和Sehumn检查了9个STR系统包括7个单个基因座分型（CSF1PO，TPOX，TH01，HPRTB，FES/FPS，vWA，F13A1）和2个复合扩增（CSF1PO/TPOX/TH01和HPRTB/ FES/FPS/vWA）分型的种属特异性。结果在检测的23种动物中有17种出现扩增产物。人类、大猩猩、黑猩猩和猩猩的DNA能扩增出9个STR系统中的8个系统。只有FES/FPS基因座引物种属特异性最稳定。但上述动物DNA材料扩增片段长度均不在人类等位基因片段大小范围之内。

英国法庭科学服务中心调查了6个STR基因座和Amelogenin性别检验基因座。非类人猿标本不能在这些STR基因座得到扩增，但是在5种动物（猪、牛、羊、狐狸和獾）中发

现了比人等位基因小 2bp 的 DNA 片段。灵长类在这些 STR 系统虽然能得到很好的扩增，尤其是黑猩猩，但是其等位基因大小通常是在所观察到的人类等位基因范围之外。

对于可能在法医物证检材中发现的 30 种微生物 DNA 模板，进行了 STR 系统的 TH01、TPOX、CSF1PO 基因座检测。在这些 STR 基因座上没有发现 PCR 产物，但是有 6 种细菌 DNA 在进行 D1S80 基因座扩增时，却出现了位于人类 D1S80 基因座等位基因大小范围内的非特异 PCR 产物。

2. 灵敏度测定

每个方法都有保证所得结果的正确性的最低检测量（即灵敏度），在其灵敏度以下的量所得到的结果不可靠，在建立新技术方法时必须进行灵敏度测定。灵敏度是与实验技术有关的指标，受诸多因素影响。一般而言，多基因座 DNA 指纹图需要 0.5μg 以上 DNA，相当于 50μl 血液；单基因座 DNA 纹印图为 25ng；PCR 用量更少，1μl 血即可。进行精斑的 DNA 指纹分析，需 1cm 直径大小的斑痕，约需 50μl 的精斑；PCR 分析只需 1～20μl 精液。分析毛发的核 DNA 需有毛囊，一根毛发的毛囊即够 PCR 分析。

3. 结果可重复性试验

用已知分型的样品评估在实验室内或不同实验室间标准方法结果的可重复性。

4. 各种条件检材试验，包括：

（1）干燥和液体样品一致性试验

用已知分型的斑迹样品与相应的液体样品验证两者是否分型一致。

（2）环境耐受检验

将已知样品暴露于不同环境，评估系统对时间和降解环境（温度、湿度、紫外线）等影响因素的反应。斑痕的时间及保存条件不同，影响也不同，新鲜的斑痕较易得到满意的结果。

（3）基质的影响检验

将体液与常见物质（如燃料、土壤）混合或涂于织物、毛皮等常见载体上进行检验。各种斑痕载体材料质地及颜色不同，对 DNA 分析影响不同，深色载体的染料在 DNA 提取过程中与 DNA 一起被提取出来，干扰酶的反应。不同标记对这些因素敏感性不同，有些分析方法分析效果较差，如一般不太容易得到血痕的 DNA 指纹图，而 PCR 为基础的标记则相对好些，尤其是 STR 基因座，几乎各种载体上的血痕均能获得理想结果。

5. 混合样品分型能力试验

检测对混合样品的分析能力，确定系统的最低量。

6. 实际案件物证检验

将新方法应用到实际案件中，检验其实用性。

第十九章 DNA分型结果的解释

科学地解释DNA检验结果，是法医DNA分析工作者的一项最重要任务。如果两个样品的相似性结果是基于不完全或错误的信息，对于检验所得到的原始遗传图谱进行评估是没有意义的。本章将讨论对DNA检验作结论时所必须要考虑的有关样品和所用的检验系统的诸多问题。这些问题有的虽曾在相关章节中讨论过，但有必要予以重复。

19.1 复杂化因素

有几个因素可能影响所有DNA分型结果，分别列举如下，同时对它们在物证意义评价中的潜在影响作简略介绍。

19.1.1 多个供体的混合样品

当一个物证样品由一个以上个体遗留时，会出现复合的DNA图谱，混合样品通常发生在像含有阴道分泌物和精液的阴道擦拭物（偶尔混有其他液体如唾液和粪便物质）混合物证中。一般衣物类有些部位易积留穿着者的体液形成的斑迹，像裤衩的档部、衬衣的腋下部位，容易产生这种情况。如果物证留存在这些部位，这些部位可能含有受害者和罪犯两个人或多个人的细胞。每个人在一个基因座有二个等位基因，那么此类含有多个人细胞的物证可以检测出多个等位基因（或带、点或峰）。在怀疑为多个人的混合斑迹情况下，一般不应人为地将两个等位基因归为一个人，尤其是所检验的样品有二个不好分离的等位基因，仅因为他们与比对样品相关而划分为一个个体。但如果性犯罪物证样品为受害人和罪犯两人的混合斑迹，受害人的分型十分明确，那么可以先根据受害人基因从物证中区分出受害人的等位基因，然后将可能是“罪犯”的基因型与比对样品比对。

19.1.2 差异裂解分步提取

阴部分泌物、口腔或直肠性攻击等物证经常出现精子和上皮细胞混合情况。这种情况下，可以通过差异裂解分步提取分离不同类型细胞，但这种分离并不总是100%成功的，精子DNA可能渗到表皮细胞部分，表皮细胞DNA可能混到精子部分，导致图谱在一个基因座有二个以上等位基因，使解释变得复杂。

而且，有的样品的表皮细胞较多而精子非常少，这种情况下，回收的精子DNA很少，在提取过程中可能会有遗留的表皮细胞，而且一些基因座的PCR扩增产物中非特异带较多，无法确定物证究竟含有几个个体，弄不清最后的分型结果究竟是来自于精液供体还是来自受害人。仔细考虑所有的分析结果，包括不同提取过程中细胞碎片的显微镜检验，将有助于正确解释结果。

19.1.3 降解

DNA在体外的任何环境下，不像在体内那么稳定，环境条件可能改变DNA的物理和化学特性，时间、温度、湿度、光照和污染（化学和生物的污染）等这些环境条件会引起DNA降解。降解DNA是否可以进行法医DNA分析，是否影响分析结果？这取决于降解程

度和所用的 DNA 分型系统。例如，RFLP 检测 DNA 片段长度大约在 0.6kb 到 20kb。如柏油路上已暴晒 2~3 天的血迹，样品中 DNA 高度降解，DNA 片段平均大小为 2 kb 范围，那么在 RFLP 分析图谱中，虽然有比这个范围小的片段存在，但不可能检测到任何大于这个范围的谱带。在这种情况下首先要考虑到的是一条带图谱并不能真正反映供体图谱的可能性，避免假排除。DNA 样品的降解程度可以通过凝胶电泳来估计，以获得平均片段大小信息，预测 RFLP 分析是否可能得到十分肯定解释的一条带图谱，或是选用 DNA 片段范围较小的遗传标记系统检验，以期成功的机率更多。

通常以 PCR 为基础的检验比 RFLP 更适合分析较小 DNA 片段，例如上述所提到的只有 2~4kb 大小的 DNA 片段完全适合目前法医学所有的 PCR 分析系统。通常，PCR 所需 DNA 最长片段不超过 1kb，STR 基因座的片段更小，即使遭受热、湿度和污染等严重条件的物证样品用 PCR 分析，其成功率仍然很高。

19.1.4 污染

物证样品主要有二种不同方式的污染，第一种是样品或载体本身固有的一些化学或生物物质，例如前面提到的裆部或腋下部分，在物证精液或血液沾在这些地方前，其他物质可能已经积累在载体上。这类物质包括生物物质如细菌、含有人 DNA 的细胞，还有像肥皂、洗涤液和除臭剂等化学物质；第二种方式是在收集及分析物证时由于不正确处理偶然将这些物质污染到物证上。不论哪一种情况，污染所造成的结果是相同的：抑制 DNA 分析；细菌分泌的酶降解人 DNA；污染有额外的人 DNA 导致错误的结果，错误地认为物证 DNA 是由多个供体混合组成等。

1. 抑制 DNA 分析

由于 RFLP 和 PCR 分析均依赖于酶（分别是 HaeⅢ、HinfⅠ和 Taq 酶）的反应。干扰酶反应的物质将会潜在抑制酶对 DNA 的作用。这些抑制剂包括平常经常遇到的各种物质，如布中的染料（尤其是斜纹布染料）和一些生物物质，尤其是血液中的血红素。尽量在 DNA 提取阶段去除这些物质，消除对酶反应的抑制。如果确定有抑制剂，可以重复某些必要的提取程序，或采取某些特殊纯化试剂盒，也可以在酶反应中加入一些中和剂来中和酶抑制剂，如 BSA（牛血清白蛋白），通常直接加到酶反应内，与酶抑制剂结合，使抑制剂在反应中失去活力。有时仍不能克服酶抑制，这种情况下，可能得不到分型结果。

2. 非人类 DNA

理论上，所有生物体含有 DNA。既然每个地方都可能因生物体的存在而留有 DNA，难免常常发现物证样品含有非人类 DNA，这些 DNA 可能来源于其他生物如动物、植物或微生物。非人 DNA 并不是一个令人棘手的问题，因为用于法医 DNA 分析的遗传标记大多是人类特异的，也就是说其它动物种类 DNA 得不到这些 DNA 分型系统的结果。但是，由于微生物能产生降解 DNA 酶的存在，易引起人 DNA 降解，使样品不能分型；而且，DNA 分型结果取决于样品中人 DNA 含量。因此分析人员必须计算出样品的 DNA 总量以及人 DNA 含量，利于正确选择标记和反应系统，尽可能得到分析结果。

19.2 不同检验系统的结果解释

前面讨论的各种因素可能互相作用影响 DNA 和随后的 DNA 分析及对结果的解释。系统地研究环境因素对 DNA 分析影响，有助于分析人员在检验中选择合适方法和程序以及对结果的正确解释。下面讨论各特定系统检验结果的判读与解释，包括具体案件中常遇到的环境条件对结果的影响。

19.2.1 RFLP

用于法医 DNA 检验的 RFLP 基因座是高度可变的，除纯合子个体外，杂合子个体是从他们的父母中遗传不同的等位基因，产生有两条带的特征图谱。由于正常的遗传变异，或由于样品本身或分析过程中人为造成的，使其与预期标准图谱可能有偏差。下面叙述一些常见的差异，讨论他们形成的原因和结果的解释。

1. 多带（多于 2 条）图谱形式

多带现象在多基因座 DNA 指纹图中不易发现，因为本来它的谱带就很多，无法检测是否存在多带，只有在与参考样品比对时，才有可能发现是否有多带。而对于单基因座的 DNA 纹印图，在 RFLP 图谱中看到二条以上的带，即表明多带的存在。不论单基因座还是多基因座图谱，关键问题是要弄清多带是一个以上个体提供的混合图谱，还是人为造成的额外带的图谱而实际样品只来自于一个个体？下面介绍单个个体而不是多个个体（混合样品）样本在分析中出现额外谱带的情况。为了便于说明，以单基因座为例。

（1）三条带图形

在某一特定基因座上一个人可能出现三条带的情况在遗传学上是罕见的（图 19－1，图 19－2），VNTR 序列内部发生点突变时，如产生一个 HaeⅢ酶切位点，则产生三条带。例如，正常表现为 8kb 带的 VNTR 将被切成 2kb 和 6kb 二个较小谱带，这种突变遵循孟德尔遗传定律（稳定传递给下一代），表明是由遗传所致，而不是一时的人为产物，由于这种情况在人群中非常稀有，因此在个人识别上非常有用，一旦比对样品与其相同，将显著增加他们来源于同一个体的可能性。

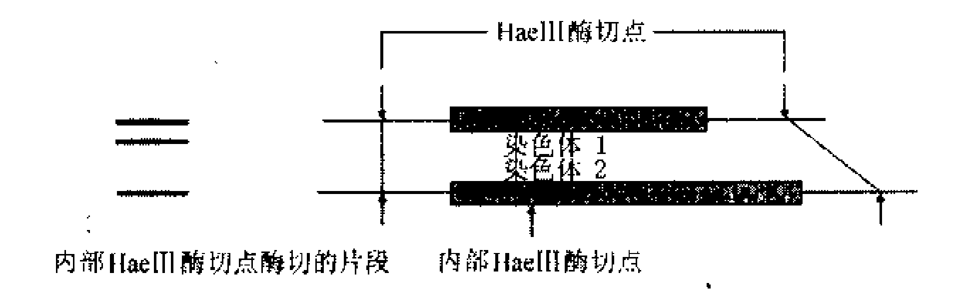


图 19－1　3 条带图谱的分子基础模式图

①限制性内切酶的部分消化（不完全酶切）

当 DNA 链上的限制性酶切位点受到化学阻碍或内切酶（HaeⅢ）的作用受到部分阻碍，或相对于 DNA 量来说酶量不足时，基因组中一些 HaeⅢ酶切位点可能没有被切开（图 19－3，图 19－4），这样有些 DNA 片段将会比完全消化后预计的片段长，这些长片段的长

度在一定程度上可以估计出来，因为重复序列酶切位点位置是固定的。部分消化的结果是一些 DNA 片段的大小正确，而另一些片段则稍微长些，既有正常的 DNA 谱带，也有看上去较大片段的 DNA 谱带，强度上一般比正常的带弱一些。通过对较大片段的大小测定，与预测的大小比较，可以确定是否有部分消化。

在大量的日常案件鉴定中对于样品是否出现有部分消化，可以先用小的琼脂糖凝胶电泳消化产物来判定是否有部分消化。对于有部分消化的样品，需要重新消化 DNA。

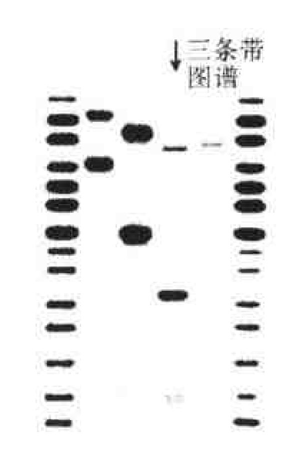


图 19-2 3 条带图谱的自显影照片

②星号活性（star activity）

酶由于星号活性，降低了特异性，酶识别比较相近

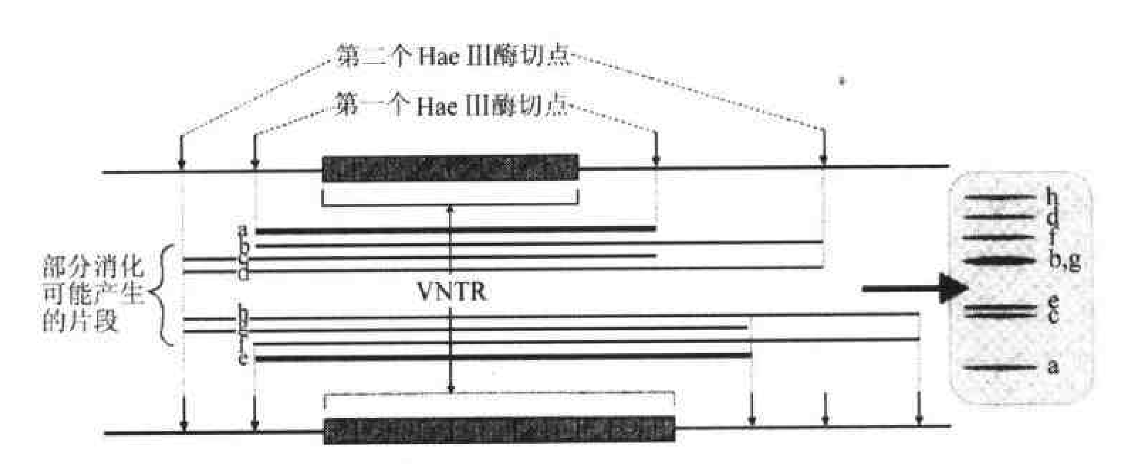


图 19-3 部分消化的分子基础模式图

序列，并在此切割 DNA，其结果是出现比真正等位基因小的谱带（图 19-5，图 19-6），导致每一个等位基因出现二条带，而不是一条带。星号活性产生的每一对片段比真正出现的谱带要小。

在实际使用中，一些分析条件包括消化后酶没有完全失活或反应在不适当的条件进行，将会导致星号活性。有时星号活性产生的带比图谱中真正的带还要强，这不仅有可能发生假排除，也可能发生潜在的假阳性的结果（尽管概率很小），对检验结果的判定是很危险的，可能造成误判。各种内切酶产生的星号活性是不同的，HaeⅢ具有有效的控制星号活性的能力。在法医 DNA 分析中选择控制星号活性能力强的酶（HaeⅢ）进行 RFLP 分析，同时使用避免引起星号活性的条件，可以避免出现这类问题。

③前一次探针没有完全洗脱掉

当杂交膜进行多个探针反复杂交时，前一个探针在下一个探针反应前（杂交前）没有被彻底洗脱掉，有时在第二个探针的杂交图谱中，还可以观察到一些浅浅的带，其位置与前一个探针分析的位置相同，在放射性同位素标记系统中更易发生这种现象。比较自显影片，观察和计算片段大小可以比较容易地判定是否为前一次探针没有被彻底洗掉所显示的带。



图 19-4　不完全消化的自显影照片

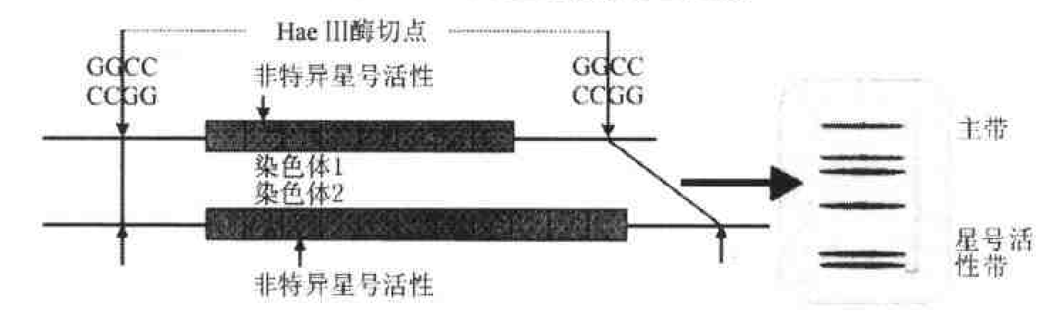


图 19-5　内切酶星号活性的分子基础模式图

④多条谱带图谱的总结

引起多条谱带图谱的可能原因有：存在多个供体，遗传突变（三条带图形），前一次杂交的探针没有彻底洗脱干净，限制性酶的部分消化，星号活性。对于造成这些多带的原因可以通过一些实验加以区分：不完全洗脱造成的多带可以通过观察比较几次杂交前后的自显影片来确定；部分消化造成的多带是出现比图谱中真正谱带大的片段。多个基因座都发生谱带比真正的谱带要大，足以证实有部分消化的存在；星号活性结果是谱带比正常的要小，一条带的强度减弱，较小片段的带强度增加，很明显表明星号活性的发生。真正的遗传三条带图谱既没有部分消化（大小与该基因座其他不符，在其他位点

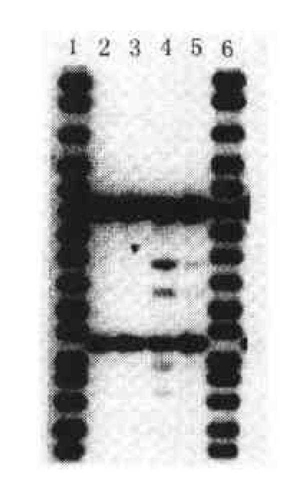


图 19-6　零星活性的自显影照片

2 号和 3 号没有星号活性，4 号和 5 号 DNA 在主带下面有浅浅的额外带，表现出星号活性

没有发现有额外带），也没有星号活性。

2. 单一条带图谱

生物物证 DNA 分析选用的 RFLP 基因座具有高度变异性，杂合度很高；检验的大多数为杂合子个体，拥有二条带，只有少数个体为纯合子，RFLP 图上只出现一条带。但是另外一些情况下，可能导致在 RFLP 图谱上只出现一条带，而事实上为杂合子个体，下面将叙述发生这种情况的可能原因。

（1）条带之间非常接近

在实际分析中，RFLP 分析胶不能分离两个在长度上仅差一个或几个重复单位的限制性片段，例如 D1S7 位点重复长度只有 9bp，而酶切片段常常在 4kb 到 12kb 范围，分析用的胶不能区分长度为 6 200 与 6 209bp 的片段，这二条带在图谱上以单一条带出现，造成该个体在该基因座表现为纯合子的假相，而事实上，在分子水平该个体为杂合子，只是由于胶的物理性质，胶的分辨率低而没有分开。这个缺点在越大的片段，变得越明显，片段大小越大，越难以分离，这种分析系统不能检测认定所有的等位基因，为此提出了一个连续等位基因系统的概念，表示邻近的等位基因之间的边缘是不断开的，就象一个等位基因刚结束，另一个等位基因起点又同时开始出现，如在前面章节中讨论的，RFLP 等位基因同一认定是以外部强加边界定义分成若干组的。需要特别强调的是，这种分组是由于现时的分离和检测技术人为造成的，而并不是有机体的遗传学所引起。

（2）在胶端的带

法医学选择 RFLP 基因座的标准之一是电脉过程中胶载留的可能谱带大小范围，常用的基因座产生的带主要在 0.6kb 到 20kb 大小之间，但是有些等位基因落在这个范围外，较大片段一般没有问题，不会电泳出胶端，他们仍能看得到，即使他们的大小不能得到测定。但是，较小的片段比较难以检测，因为在胶的前沿，扩散严重，在电脉过程中，最小的部分（尽管很少）确实被泳动出去了，而留下单条带图谱，看上去像个纯合个体，但事实上在这个给定的基因座是杂合子，不能检测出较小等位基因也是分离和检测技术上的人为结果。

（3）降解

DNA 降解到一定程度，片段长度小于一定片段长度的检测范围或已没有足够量的大分子 DNA，以致在这个区域内观察不到 RFLP 图谱（图 19－7，图 19－8，图 19－9）。造成 DNA 降解的因素强度愈大，时间愈长，DNA 片段的平均大小变得越小。通过琼脂糖凝胶来检测降解，并与在同一胶中的标准分子量比较，估计 DNA 片段平均大小，根据结果可确定是否有降解存在。对于降解比较严重的样品，一般不排除也不肯定的结论是明智的。比如说对拥有二条带分别为 10kb 和 4kb 的杂合个体来说，如分析样品的平均大小只有 7kb，显然，10kb 的带就检测不到，而只能检测到 4kb 的带。这种本应出现二条带杂合子个体样品检出一条带的原因是由于外部环境等因素造成的 DNA 降解所致。

降解导致的单条带图谱现象常发生在物证样品，而比对样品较少发生。因为在提取物证前无法控制物证样品的降解，而比对样品是专门为 DNA 分析要求而采集的，比较新鲜、采取了正确的采集、保存和送检方法，防止了降解发生。大多数情况下，在图谱中，比对样品有二条带，而物证样品只有与比对样品片段大小相同的较小的一条带。如果比对样品

图 19－7　降解的分子基础模式图

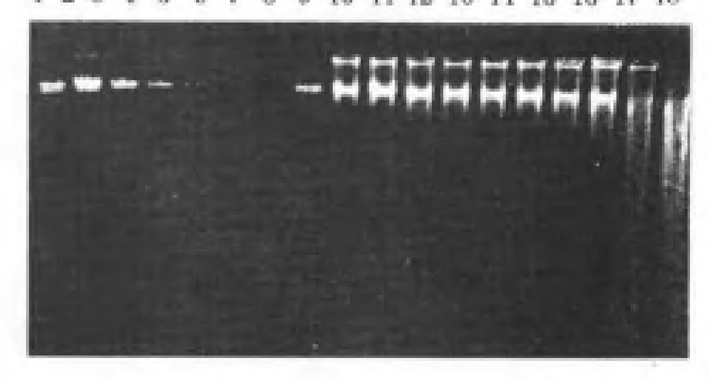


图 19－8　琼脂糖凝胶检测降解程度

1－8 号为标准和对照 DNA，9－18 号为 65℃处理不同时间的精斑 DNA，17 和 18 号为同一 DNA 处理 1 个月，DNA 已发生降解，探针的自显影照片为图 19－9。

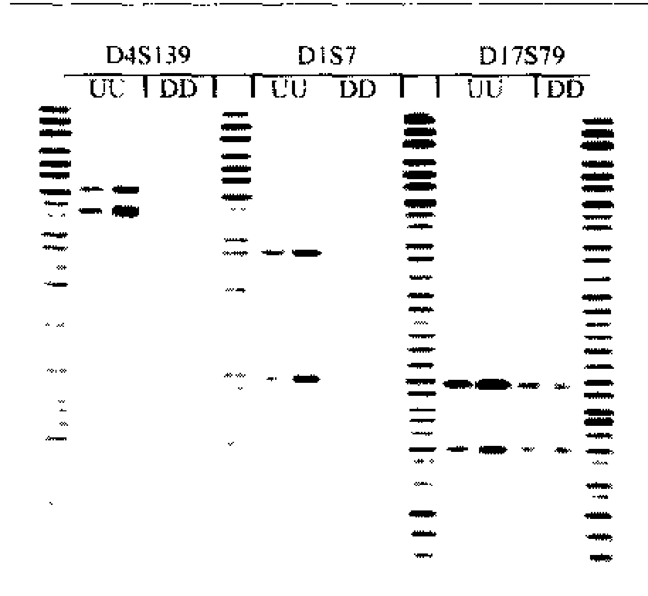


图 19-9 降解 DNA 自显影照片

U、U 为图 19-9 中的 15 和 16、D、D 为图 19-9 中 17 和 18 号，降解非常严重，在 D4S139 基因座已经没有带，在 D1S7 基因座只留下小片段，D17S79 基因座，由于探针检测的片段较小，还可以得到 DNA 谱带。

采自不新鲜尸体的血样或腐败组织，或采集后没有适当保存和送检，尸体样品 DNA 有严重降解，而从嫌疑人衬衫上发现的干血痕物证比来自尸体的比对样品保存得好，则物证样品会出现二条带，而比对的尸体血样只有一条相同的小片段带。

(4) 单条带图谱的总结

有四个原因影响任一 RFLP 基因座单条带图谱的出现：真正的纯合子；两条带之间挨得很近；极其小的谱带在电泳分离时被泳出胶外而没有被检出；DNA 发生降解，破坏了二个等位基因的较大等位基因。后三种情况中杂合子个体的图谱以纯合子单条带方式出现，有经验的分析人员会非常谨慎、科学地解释这种结果。

3. RFLP 解释问题的总结

这些高变 RFLP 基因座的高度个体差异性（区别个体间能力）使其在法医学检验中一度得到了极为广泛的应用。但是物证检材 DNA 因环境因素等造成的降解，必使 RFLP 的应用受到了限制，使其结果解释复杂化。对于每一案件的检验结果都要把这些因素考虑进去，判断哪些是人为情况，以避免造成严重的结论错误。了解每一方法的优缺点，来解释样品的特殊分型，区分真正的图谱与人为的图谱。

19.2.2 PCR 系统

PCR 扩增长度多态性的系统，以电泳方式检测其多态性。PCR 产物在胶上电泳分离，以类似于 RFLP 方式阅读结果。基于等位基因特异序列差异检验特定序列是否存在（有信号）或不存在（没有信号）的标记系统，其检测方式则通常是以斑点或滤膜条杂交方式，每一个点或条表示一个等位基因的方式来阅读结果。但以 PCR 检测的基因座多态性比 RFLP 低得多，二个染色体拥有相同等位基因（纯合子）的概率要比 RFLP 多。

法医物证鉴定中早期应用的 PCR 系统，每个基因座都是独立扩增，随后进行检测，后来技术日趋完善，多个基因座可以同时在一反应管中进行（复合），扩大了 PCR 的应

用。同样量的 DNA 可以进行三个或更多个基因座的分析，复合扩增基因座可以在同一胶或滤条上同时分析。

但不管系统之间有无差异，有一些问题是所有 PCR 反应共有的，下面先介绍这些共同性的问题，随后就标记特有的问题进行讨论。

1. PCR 一般问题

（1）混合样品

PCR 系统中的主要问题是样品含有一个或二个以上个体的 DNA，等位基因不能均等地得到平衡扩增，致使有些带或点强度不同。研究表明样品按一定比例混合，当混合比例达到足够悬殊时，较少量 DNA 的等位基因扩增会受到抑制而达不到检测水平，例如，当样品比例大于 1:20 到 1:50 时，则不能检出少量部分样品的等位基因，此时样品不表现为混合样品，如果两个样品来自于二个不同的来源，而且二者比例接近（大约 1:10 到 1:2），则可明显表现出来自一个体的等位基因强度比另一个体的强度要强一些（深些），当强度差异达到足以区分的水平时，有经验的分析人员就可以区分出，并且进行分类命名各自的基因型。

（2）扩增参数

PCR 反应中的每一成分都是必需的，在最佳反应条件（反应物浓度、特定变性、退火、延伸温度及各温度下的反应时间和扩增循环数）下进行扩增，保证系统不发生优先扩增等问题。

（3）优先扩增

优先扩增指的是在同一反应中，一个等位基因比另一个等位基因得到更多的扩增，结果可能导致杂合子（两个不同等位基因）中两个等位基因带强度（峰高）差异显著，甚至表现为纯合子。这种现象也称为等位基因的流失，引起优先扩增现象有以下几个原因：

①随机波动　由于加人 PCR 反应中的 DNA 量极少，杂合子个体中两个等位基因在 PCR 前几轮循环中彼此可能发生不均等扩增，在扩增过程初期出现一等位基因量比另一等位基因的量多很多，随着扩增，这种状况将更加加剧，较多的那个等位基因优先扩增，而另一个等位基因则仍然滞后扩增，最终结果是只检测到二个等位基因中的一个，导致错误结论——样品来自于纯合子个体，而不是杂合子。因此，必须了解方法的灵敏度和样品的 DNA 含量。

②等位基因片段长度范围较宽　Taq DNA 聚合酶的一个缺点是：短小等位基因比大等位基因扩增效率高，这对于序列多态性标记没有影响，但对于长度多态性标记如 D1S80 基因座就有影响，必须注意这个问题。小卫星基因座的等位基因重复序列较长，等位基因片段大小差异大，会出现较大等位基因片段得不到全长扩增，拷贝数少，带的强度减弱，造成两个等位基因强度有差异，易被误认为样品为含有多个个体的混合样品，或根本检测不到较大等位基因片段，将本为杂合子的图谱误认为纯合子。解决这个问题的办法是在 PCR 反应液中加 BSA，BSA 可以使等位基因的强度趋于平衡，如图 19 - 10。

STR 基因座由于重复序列较短，等位基因数较少，较小等位基因片段与较大等位基因片段之间差异较小，较小等位基因优先扩增现象不严重。

③等位基因序列差异　G - C 的化学键比 A - T 的强，因此 G - C 含量高的 DNA 片段要求更高的温度使之变性。如果一个等位基因的 G - C 含量比平均的高，那么在较低的温

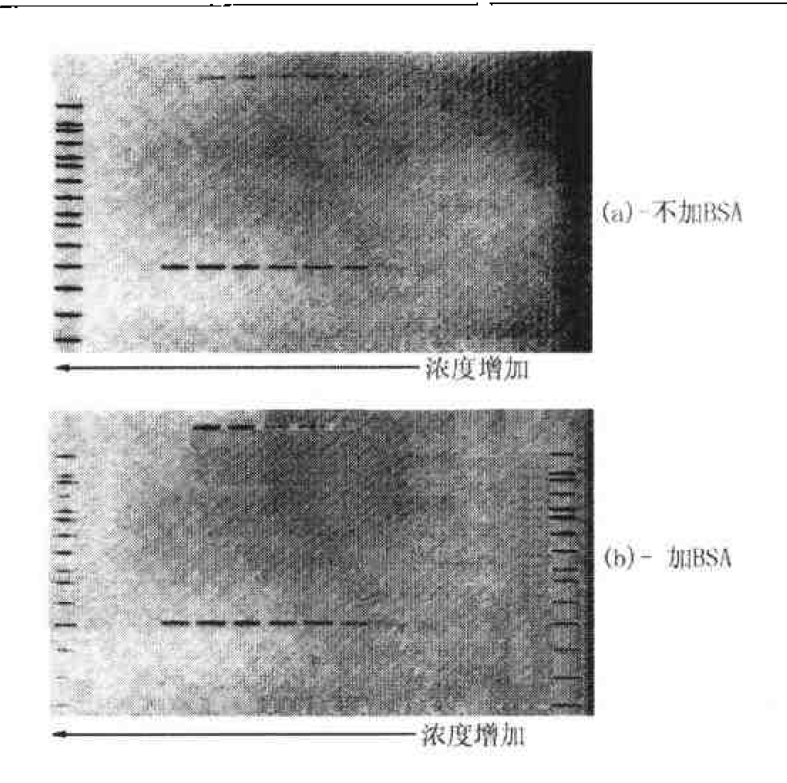


图 19-10 BSA 与优先扩增

在 D1S80 基因座含有两个等位基因的 DNA 样品，DNA 浓度从右到左增加。a. 不加 BSA 的 D1S80 基因座银染图谱，尤其在高浓度下，较小等位基因优先扩增。b. 反应液中加 BSA，等位基因较为平衡，优先扩增效应减缓

度下有可能不能完全变性，妨碍随后的引物结合，使 G-C 含量高的等位基因扩增效率降低。严格控制每一步骤的温度以克服这种现象。

④异二倍体（heteroduplex） 核苷酸序列相近但又不完全相同的 DNA 两条链退火形成异二倍体，对于大多数 PCR 反应，退火条件是严格特定的，一般不会发生异二倍体形成，但对于一些十分相近的序列，如 HLA DQα/A1 系统中等位基因 1.3 和 3，有可能形成异二倍体。

⑤酶的抑制 PCR 反应依赖于 DNA 聚合酶，每个酶都有其最佳反应条件。Taq DNA 聚合酶对反应条件较为敏感。在模板 DNA 提取过程中与 DNA 一起被提取出来的一些物质如血红素、染料会抑制 Taq DNA 聚合酶。解决的方法有三条：一是增加 Taq DNA 聚合酶量；二是减少模板 DNA 量以稀释抑制物；三是加 BSA。实验证明每一种方法在不同条件下均有效。添加 BSA 方法是最常用的方法，一些实验室和一些商品化的试剂盒内都加 BSA。

⑥阈值检测对照 在扩增及检测过程中设置一个能扩增、检测最低量的阳性对照是十分必要的，这有助于确定一些强度弱或浅的带是真正的等位基因，还是一些人为造成的带。HLA DQα/A1，Polymarker 等一些系统都设置一个内部阳性对照。

2. 序列多态性 PCR 标记系统

序列多态性分析方法很多，大多采用寡核苷酸探针杂交检测 PCR 产物方法，已有大量试剂盒供应，它们采用反相杂交方式检测，如 HLA DQα，Polymarker 等，结果非常简单

——"有"与"无"，即"0"和"1"方式，直接确定结果。但是必须清楚反相杂交结果阅读原则与技术本身的固有限制：(1) 无法正确判定含有多个供体样品的型别；(2) 探针间的交叉杂交；(3) 差异扩增与降解。

3. 长度多态性 PCR 标记系统

(1) 小卫星基因座

小卫星基因座利用电泳分离检测等位基因，以重复单位数目命名，根据同一凝胶中等位基因分型标准物 (LADDER) 确定基因型。对于 LADDER 以外的等位基因，如 D1S80 基因座，LADDER 中等位基因为 14～41，对于大于 41 的等位基因，这些带很难正确确定他们的重复单位数，把他们归类为">等位基因 41"。

①优先扩增　由于小卫星基因座的重复单位序列较大，等位基因大小差异较大，易造成较小等位基因优先扩增，较大等位基因丢失现象，杂合子个体表现为假纯合子。这种现象可以通过在 PCR 反应液中加入牛血清白蛋白 (BSA) 得到缓解 (图 19－10)。

②降解　另一个易造成分型错误的是 DNA 降解。DNA 降解引起的结果如同 RFLP 系统，较大等位基因丢失，只剩下较小等位基因。

③异二倍体形成　异二倍体形成并不是 D1S80 基因座主要问题，但异二倍体的形成可能产生人为带，在扩增产物检测凝胶中出现一些浅带。这种现象一般发生在模板 DNA 过量的情况下。

④阴影带　偶尔在 D1S80 等位基因主带下面出现一条强度较弱的带，称之为"阴影带"(shadow bands)，阴影带产生的机制还不太清楚，有人认为是在扩增过程中酶的滑动或/和电泳过程中的凝胶条件造成的，阴影带在分型过程中很少作为一个问题，因为他们很容易确定为人为的杂带。

(2) STR 系统

①双带　为了更好地分辨每一个 STR 等位基因，扩增产物用变性聚丙烯酰胺凝胶系统分离。片段完全互补的两条链由于碱基组成的差异，在凝胶电泳中的迁移率不同，结果较小的等位基因分辨成双带 (图 19－11)，电泳后染色，一个等位基因的两条带被均匀地染色。样品和等位基因 ladder 都出现这种现象，重复性较好，所以在解释时不会成为一个问题。

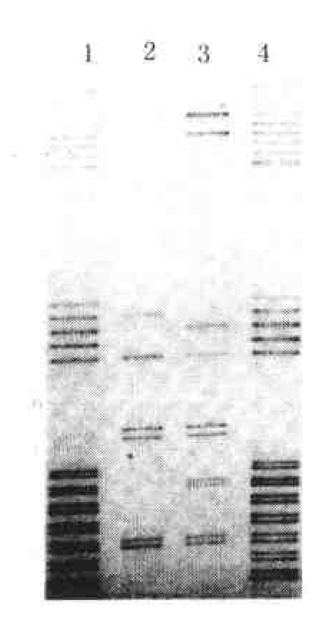


图 19－11　STR 银染结果——双带现象

荧光标记一条引物的链，使扩增产物中只有一条扩增产物带荧光标记，检测时只观察到一条带，结果更容易理解。

②Stutter 带　Stutter 带是指出现在比其主带 (等位基因) 短一个重复位置的带 (如图 19－12)，原因可能是在复制过程中 Taq DNA 聚合酶的滑动。一些 STR 基因座易出现 Stutter 带，这与序列有关，复制滑动在二核苷酸 STR 基因座多见；四、五核苷酸重复序列基因座则相对少见。在复合扩增中，由于扩增反应条件不是每个基因座的最佳条件，容易出现 Stutter 带。较大等位基因的

Stutter 带百分比要比较小等位基因的大，例如在同一基因座，等位基因 10 比等位基因 3 更易出现 Stutter 带。Stutter 带在混合样品分析时是一个问题，因为有时很难区分或确定一条带是 Stutter 带还是来自另一个体 DNA 的真正等位基因。Stutter 带的阈值定为主带的 5%～15%。如果小于主带的 15%，则认为是 Stutter 带；如果超过了主带的 15%，则解释为真正的等位基因。

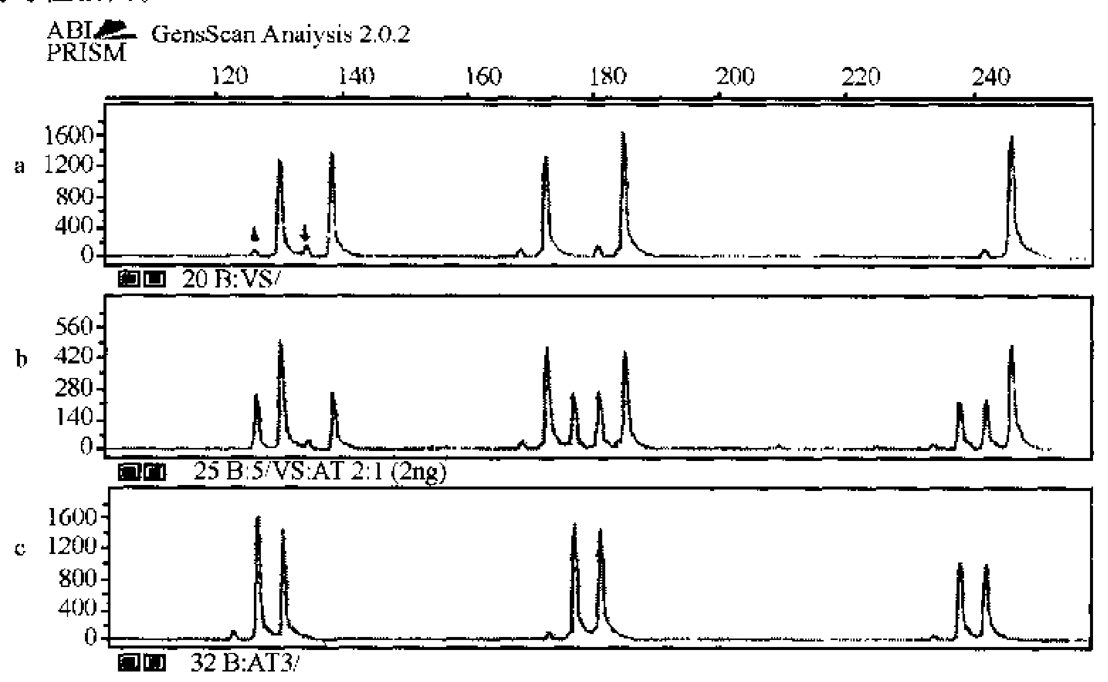


图 19－12 STR 基因座的 STUTER 峰与混合样品

(a) 和 (c) 为两个单一的个体，箭头所指为 STUUTER 峰；

b 为 (a) 和 (c) 中两个个体按 2a∶1c 混合成的混合样品结果

③pull－up 峰 多色荧光标记检测系统中，由于有机荧光染料发射光谱较宽，不同荧光之间相互有干扰，产生干扰峰，称之为 pull－up 峰。不过这种峰的峰形与真正的等位基因峰不同，而且出现的位置与另一颜色的峰位置一致，较容易区分开。

④非产物荧光污染 一些具有荧光的物质如抗生素、维生素、多环芳香物、荧光素发生体和染料等在电泳图中出现与 DNA 荧光片段相同的峰，可能会干扰 DNA 分型，影响鉴定结果。Chelex DNA 提取方法甚至有机提取法不能去除抗生素、维生素、多环芳香物、荧光素发生体、各种染料等，以 Chelex 提取法提取的 DNA 为模板的 PCR 产物常伴有荧光污染峰。不过，干扰峰通常较宽，比较容易与荧光素染料标记的 PCR 产物峰区分开来。

非产物荧光污染产生有以下四种可能的情况：一是带色载体上的体液斑，在提取过程中染料被浸了下来；二是植物体上的体液斑，叶绿素可能与 DNA 同时提取出来；三是病理状态下个体的血或组织样品，如铅中毒、某些形式的卟啉症，此时血中的卟啉浓度大幅度提高；四是年轻时曾接受过环组抗生素治疗的个体的牙或骨组织（成长的骨和牙齿吸收、蓄积了这些抗生素）。采用空白基质（载体）对照实验，或将未扩增的物质与 PCR 扩增产物一起电泳检测进行比较，确定是否有非产物污染峰。

⑤样品为混合样品 当分析样品为混合样品时，经常出现多带现象，这是造成额外带

的最常见原因。混合样品情况将在后面详细介绍。

⑥电泳条件不适当　FGA基因座的带会移到vWA基因座区域，造成vWA基因座有三条带，而FGA基因座只有一条带，重新电泳检测可克服。

⑦一些复合扩增体系中，同一颜色的基因座间等位基因片段有些交叉，主要是一些片段较大的或较小的稀有碱基位于某一基因座的等位基因内部。如FGA的稀有等位基因14位于vWA基因座。

⑧模板DNA极微量、DNA降解、有抑制剂存在　模板DNA极微量或DNA降解或有抑制剂存在，在PCR开始几个循环，引物退火的随机差异，导致某一个等位基因优先扩增。另外，当一对等位基因的大小差异很显著时，也可引起一个等位基因的优先扩增。

如果一基因座二个峰高比值小于70%，而没有其他指标表明是混合样品，那么该样品必须重新扩增以确定不平衡情况是否有重复性。不平衡扩增提示标本可能为混合样品。

模板DNA极微量或DNA降解或有抑制剂存在，不同基因座扩增效率不同，有些基因座不能得到扩增，片段较大的基因座最先受到影响，如D7S820，D18S51，FGA基因座。

⑨引物结合位置的序列差异　模板DNA的引物结合位置出现突变，可能直接影响退火效率，尤其是突变发生在引物的3′端，延伸可完全被阻断，使某个等位基因无扩增产物，成为无效等位基因（null allele），杂合子个体只有一个基因产物，成为所谓假纯合子的现象，如MBP-B、D19S253及Profile plus中D8S1179基因座均可出现上述无效等位基因，不过这种情况造成的假纯合子的等位基因峰高比真正纯合子的低许多，由此来确定是否存在无效等位基因。如果突变发生在引物结合的其他位置，引物退火温度受到影响，导致信号降低。

没有一对引物可以绝对地说没有无效等位基因现象。引物位置点突变发生的机率极小（大约0.01~0.0001基因座），一个个体不可能在二个基因座受到影响。改变引物序列，避开前一引物退火区域，可以克服无效等位基因的问题。

不管怎样，只要用同一套引物相同条件分析现场物证与比对样品，如果现场物证分型与嫌疑人的相同，说明它们来自于同一来源，无须考虑无效等位基因问题。这种情况只适合于一个实验室内。如果建立数据库，实现数据共享，进行联网比对，潜在的等位基因丢失可能导致假排除。为了克服这个潜在的问题，需要降低数据库搜索的标准，允许可以有一个基因座不完全匹配，而不是所有基因座必须完全匹配。但无效等位基因对于亲子鉴定，存在假排除的可能，要引起注意。

⑩样品为混合样品　在实际检验中，明显干扰峰面积平衡的常见原因是混合样品，而不是等位基因的优先扩增。

当分析样品为混合样品时，混合样品组分的混合比例不同或混合的比例相同但拥有一个共同的等位基因，可能会发生这种基因座内的优先扩增现象，等位基因峰面积明显不平衡。扩增分析多个基因座，综合分析整个图谱，确定是否为混合样品，有助于结果的解释。

⑪微变异体与稀有等位基因　由于碱基插入、缺失或变化造成等位基因序列差异，一些重复单位碱基数发生变化，等位基因内部有不完整的重复单位存在，这些微变异体与完整重复单位的等位基因差异极其微小，而且比较稀有，不包含在分型标准物中，与分型标

准物中等位基因只差1到2个碱基。

D18S51，D21S11 和 FGA 基因座出现微变异体较多，“.1”、“.2”、“.3”，很难辨认，很难进行正确命名。如果命名不正确，会导致结果错误，尤其是进入数据库进行比对。要求认真检查分子量内标，确认等位基因的片段大小，证实确为稀有等位基因，进行正确命名。

一些片段较大或较小的等位基因，可能落在该基因座等位基因分型标准物两侧以外，在排除“pull - up”峰和非特异性扩增产物外，必须将此峰作为真正的等位基因来处理，按国际命名标准进行命名。

⑫部分 X 和 Y 染色体序列表现出高度的序列相似性　一些 Y - STR 基因座即使采用非常特异的扩增条件，女性 DNA 仍能在 Y - STR 基因座得到扩增产物，表观表现为二个男性 DNA 混合的样品。因此，对于这些基因座，要求用案件涉及的女性样品作对照同时进行分析，以确认是否有女性的干扰。

19.3 线粒体 DNA 测序

一个细胞内线粒体 DNA 的高拷贝数，使之分析灵敏度提高。一些用染色体遗传标记分析不能得到有效结果的检材，有时可成功地得到 mtDNA 分析结果。

但要注意污染问题以及线粒体 DNA 存在的异质性。同一个体的不同组织、器官或同一组织的不同部位，存在一个碱基的差异，使两者的 DNA 序列不同，异质体一般一个个体发生在一个碱基处，极少观察到在两个碱基位置同时存在异质性。两个样品在一个碱基位置的不同，不能轻易排除它们的相关性。

第二十章　个体识别

20.1 个体识别的基础与原理

法医 DNA 分析的最大一个领域是对生物物证进行个体识别，进行同一认定。DNA 分型的整个目的是为了弄清物证样品与某一个体的相关样品是否有联系，是为了去证实生物物证来源于所调查的某一特定个体的假设，为案件侦查提供线索，为审判提供科学证据。

除同卵双生子外，每个生物个体的 DNA 分子是不同的，独一无二的；DNA 遗传标记是终生不变的，不随营养、环境条件而改变。这两点是 DNA 遗传标记能进行法医学个体识别、同一认定的基础。

个体识别的原理是：在保证受检验的样品组织无突变（如癌变）、结果正确前提下，如果比对图谱的分型结果不同，可以排除两个 DNA 来自于同一个体；如分型相同，不能排除他们来自于同一个体，这有两种可能情况，一是样品确实来自于同一个体，二是样品来自于不同个体，只是所检验的标记分型一致，随着分析的遗传标记的增加，他们的图谱会不同。此时需要增加 DNA 标记，当检验的联合偶合率在 10×10^{-12}以下时，可以认定它们来自于同一来源。因为此时的偶合率表明，只有在千亿人中才会有可能出现两个无关个体随机相同，而如今地球人口不足百亿，按概率推算，不可能来自二个无关个体。可见，排除是绝对的，认定是相对的。

20.2 个体识别的遗传标记

几乎所有的多态性遗传标记都可用于个体识别，包括 RFLP、VNTR－PCR、STR、线粒体 DNA 以及各种序列多态性标记，但每个标记的识别率不同。因而，为了达到个人同一认定，需要选择合适的一定数目的遗传标记。

DNA 指纹图具有极高的识别率，实验数据表明，一个多基因座探针的 DNA 指纹图的偶合几率在 10^{-12}以下，识别率可达到个人同一认定水平，除同卵双生子外，没有二个个体的图谱相同。如果图谱不同，可以直接排除，如果相同，就可以认定为同一个体或来源。

单基因座 DNA 纹印图的识别率也较高，三个探针的 DNA 图谱可以达到个人同一认定，为了避免因某种原因可能造成的错误认定，一般要求 5 个单基因座探针的图谱相同，才能认定来源于同一个体。

对于扩增片段长度多态性系统，由于每个遗传标记的等位基因数目有限，个体识别率相对较低，要达到个体同一认定，需要分析较多的遗传标记。美国建立的 CODIS 系统选定的 13 个 STR 基因座在 7 个群体中的随机个体匹配概率如表 20－1 所示，这 13 个基因座的平均匹配概率小于一万亿分之一，除同卵双生子外，没有两个个体拥有相同的图谱，可以达到个人同一认定。

表 20-1 7 个群体中 CODIS 系统 13 个 STR 基因座的随机个体匹配概率估计

	人群						
	牙买加人	非裔美国人	欧裔美国人	瑞士人	意大利人	中国人	阿帕切人
平均 $\theta=0$	9.50×10^{14}	1.32×10^{15}	8.23×10^{14}	1.25×10^{15}	7.45×10^{14}	4.85×10^{14}	1.77×10^{12}
$\theta=0.01$	1.04×10^{15}	1.46×10^{15}	9.12×10^{14}	1.39×10^{15}	8.22×10^{14}	45×10^{14}	1.88×10^{12}
最常见纯合子（所有基因座）							
$\theta=0$	1.61×10^{13}	2.48×10^{13}	7.40×10^{13}	1.61×10^{14}	1.38×10^{14}	2.71×10^{13}	2.27×10^{10}
Upper 95% CI	5.18×10^{12}	7.36×10^{12}	2.15×10^{13}	4.67×10^{13}	4.29×10^{13}	5.10×10^{12}	8.77×10^{9}
$\theta=0.01$	1.20×10^{13}	1.82×10^{13}	5.32×10^{13}	1.14×10^{14}	9.85×10^{13}	2.00×10^{13}	1.84×10^{10}
Upper 95% CI)	3.91×10^{12}	5.52×10^{12}	1.58×10^{13}	3.38×10^{13}	3.13×10^{13}	3.85×10^{12}	7.21×10^{9}
条件概率（conditional probability）							
($\theta=0.01$)	3.82×10^{12}	5.65×10^{12}	1.52×10^{13}	3.10×10^{13}	2.75×10^{13}	6.10×10^{12}	8.26×10^{9}
Upper 95% CI	1.30×10^{12}	1.79×10^{12}	4.71×10^{12}	9.58×10^{12}	9.09×10^{12}	1.26×10^{12}	3.34×10^{9}
最常见杂合子（所有基因座）							
$\theta=0$	9.20×10^{10}	2.16×10^{11}	2.01×10^{11}	1.79×10^{11}	1.03×10^{11}	5.86×10^{10}	5.09×10^{9}
Upper 95% CI	4.63×10^{10}	1.02×10^{11}	9.64×10^{10}	1.27×10^{11}	5.29×10^{10}	2.29×10^{10}	2.65×10^{9}
条件概率（conditional probability）							
($\theta=0.01$)	6.52×10^{10}	1.48×10^{11}	1.37×10^{11}	1.79×10^{11}	7.27×10^{10}	4.32×10^{10}	3.85×10^{9}
Upper 95% CI	3.39×10^{10}	7.27×10^{10}	6.82×10^{10}	8.92×10^{10}	3.87×10^{10}	1.73×10^{10}	2.07×10^{9}

注：匹配概率为人群人数（N）分之一，N 为表中的数

20.3 个体识别

20.3.1 个体识别的三种可能的结果

物证样品（如生物体液或组织）与比对样品（典型的为血液样品）进行一系列 DNA 检验，检验完毕后，分析人员根据检验结果可以对物证样品与比对样品间的遗传相似性作出决定，有三种可能结果：

1. 分型不同，表明样品来源于不同个体，即排除。排除的结论对个人识别是绝对的，不需要更进一步讨论分析。

2. 检验结果不太肯定样品是否有相似的分型，也就是说检验所得到的图谱不太清楚，无法判读，不能确定（Inconclusive）。这种情况的发生有多种原因，包括样品降解、污染或试剂的失效（例如限制性酶的抑制剂）。这种情况下，可以用已提取的样品 DNA 或重新取材（包括从物证的不同部位）提取 DNA，重新重复一遍分析过程，以期得到明确的分型结果。当分析后还如此，没有比重新分析前获得更多的遗传信息，我们就可以认为结果不能确定。不能确定的结果没有多大意义，不需要进一步分析。

3. 分型相似。确定样品有相似性，表明可能来源于同一来源。需要进一步分析和讨论。

20.3.2 相似性意义

1. 相似性

物证与有关样品之间的相似性，经常用匹配（match）或一致或相同这些词来描述。法医 DNA 分析工作者对匹配，用这个词极窄的含意。匹配，只是表明二个样品在进行的特定分析中没有观察到显著的差异而已，并不表示绝对的匹配，可能二个样品来源于同一个体，也可能分别来源于两个个体。因为 DNA 检验只是检验了人基因组相对比较小的一

部分，只是目前所用检验方法没有检出这二个样品间存在的差异，进一步分析可能会发现他们之间存在差异，得到不同的结论。事实上，遗传相似性，或一致性的结论只是描述这样一个事实——在已进行的分析检验中，二个样品之间没有发现差异。由于 DNA 标记具有极高的个体识别能力，经多个高识别率 DNA 基因座检验后表现出遗传一致性的样品，可以达到个人的特征，认为是同一的，来源于同一来源。

（1）连续等位基因体系

RFLP 标记（DNA 指纹图和纹印图）被认为是连续等位基因系统，无法区分开离散的单个等位基因，图谱中的一条带可能含多个等位基因。比如一个等位基因含有 99 个 10bp 长的重复单位，很难与含有 100 个这样重复单位的等位基因分开。因为用于 RFLP 中的电泳系统分辨率低，不能分开 990bp 与 1 000bp 片段。因此，肉眼观察后，需借助计算机来算出片段大小。

比对 RFLP 体系图谱时，第一步是用肉眼检查图谱，根据经验和知识，分析人员确定图谱间是否有相似性，以便确定是否需要进一步分析。如果图谱明显不同，不需要进一步测定片段大小来确认不相同；如果二个图谱中的谱带只是轮廓相似，其他仍不相似，不足以下相似的结论。此时，需要借助计算机测量片段大小。

计算机测定的片段长度存在因样品、特定实验室方法学和测量系统造成的系统偏差，描述遗传相似性意义时必须考虑这种分析方法测量的准确度。当二个样品说是匹配时，就意味它们在系统偏差内是相同的，尽管有可能它们的大小不是真正相同。实验分析、比较来自于同一个体的物证和参照样品的片段长度时，发现碱基数目上会有些小的差异，计算机把它当做另一条带计算，这些差异是由于方法的差异及样品与样品的差异造成的。实验已经证实有量的差异性，即来源同一个体的二个样品由于 DNA 量的不同造成片段大小的差异。因此在确定相似性时需要一个匹配标准（match criterion）。匹配标准，通常用谱带大小的百分比表示，其含义是指二条带的测量长度只要落在彼此一定的测量范围内，可以认为是相同的，比如，匹配标准为 ±2.5%，那么二个样品的谱带大小必须在它们平均大小的 +2.5% ~ −2.5% 内，才能认定他们具有相似性。

（2）不连续（离散）等位基因系统

PCR 系统中，每个基因座等位基因较少，等位基因也较小，比较容易被分离，标记总体上为离散的等位基因，可以明确区分开等位基因。二个样品的基因型要么相同，要么不相同，模棱两可的现象只在像混合样品或样品较少不能得到完全的基因型的情况下出现。

2. 相似性意义

DNA 图谱出现遗传相似性有下列三种可能情况：

（1）样品属于同一来源，这意味物证样品，如血痕、精液、唾液等与比对样品来自于同一个体。

（2）相似性是一种重合，即是比对样品个体与物证真正的遗留者在所分析的特定标记上拥有相同的遗传图谱，然而物证样品真正来自于比对样品以外的其他个体，而不是比对样品的提供者。

（3）相似性是一种偶然，是由错误造成的。物证样品并不真正来自于比对样品个体，而是由于在收集、分析及操作者的错误导致物证和比对样品出现相同的 DNA 图谱。

物证的检验分析人员、案件的侦查、调查和审判人员以及其他办案人员，都要弄清所得到的 DNA 分析鉴定结果究竟是上述三种情况种的哪一种，以确定其对案件的意义。在物证搜集和所有的分析过程中实现严格的质量控制，可以杜绝上述第三种情况的发生。对于第一和第二种情况，要看他们的相似性大小。

20.3.3 **偶合率和似然比计算**

1. 偶合率

一旦物证样品和比对样品之间确定有一种相似性（匹配 DNA 图谱），这种相似性究竟有多大就成为下一个关键问题。相似性的大小取决于拥有相同遗传图谱每一种可能情况的各自力度。如果人群中许多个体有这种分型，那么这种相关意义就很少，因为它意味着人群中随机个体均有可能拥有相同的这种分型。反之，如果除了比对样品外，其他个体拥有这种分型的可能性很低的话，那么这种相关性意义就很大，预示着有可能来于同一个体。

通常用偶合率表示相似性大小。也就是说，除比对样品个体外，其他人是物证真正的遗留者时，其比对样品与物证图谱匹配的可能性有多大，或者随机匹配的可能性或偶合率是多少？偶合率，即在相关人群中，这种图谱出现的概率。分析某标记在某人群中有代表数量的个体，观察每一种基因型出现的次数，计算基因型频率，获得偶合率，对于只有少量等位基因的遗传标记系统，调查一定量的个体，完全有可能重复观察所有基因型。

对于有较多等位基因，问题就变得复杂多了，对于象 RFLP 基因座，一些基因座有 50 或更多个等位基因，如大约 1275 个基因型，4 个这样的基因座联合等位基因的结果有(1275) × (12175) × (1275) × (12175)，即约 2.6×10^{12} 个可能的组合基因型，分析人群的每一个体获得所有分型，显然是不可能的。另外，地球上只有 50 亿人口，大部分组合基因型并不存在。

根据群体遗传理论，对于处于 HW 平衡的群体，HW 平衡和连琐平衡定律允许依据等位基因频率来估算理论的基因型频率，纯合子个体基因型频率为该等位基因频率的平方，即 p_i^2；杂合子则为两等位基因频率的乘积的 2 倍，即 $2 \times p_i p_j$，以这估算的基因型频率理论值来计算偶合率。因此排除是绝对的，认定来自于同一个体则是相对的。相似性的意义取决于他们相似性的大小。

偶合率大小与分型系统、标记的等位基因在人群中分布情况以及基因型相关。

偶合率计算公式：

$$P_m = P_r(A_iA_j) = \begin{cases} p_i^2 + \theta p_i(1-p) & \text{当 } i=j \text{ 时} \\ 2p_ip_j & \text{当 } i \neq j \text{ 时} \end{cases}$$

式中：θ 为群体亚结构系数，p_i、p_j 为等位基因 i 和 j 的等位基因频率。

当比对样品有亲属关系时，偶合率计算公式为：

$$P_m = Pr(A_iA_j) = \begin{cases} \varphi_0 p_i^2 + \varphi_1 p_i + \varphi_2 & \text{当 } i=j \text{ 时} \\ 2\varphi_0 p_i p_j + 1/2\varphi_1(p_i + p_j) + \varphi_2 & \text{当 } i \neq j \text{ 时} \end{cases}$$

φ_0、φ_1 和 φ_2 分别为后代有 0、1 和 2 个等位基因相同（identical by descent，IBD）时的概率。当为同胞时，(φ_0、φ_1 和 φ_2) = (1/4，1/2，1/4)；当为第一代堂婊亲属，(φ_0、φ_1 和 φ_2) = (3/4，1/4，0)

n 个独立遗传的标记的联合偶合率计算公式：

因此，检验时不能单凭一个碱基序列的差异轻易排除其相似性，3个以上碱基不同排除较为可靠。同一母系人员间序列也存在异质性，在亲缘鉴定时要予以注意。英国的 FSS 和美国的 FBI 及军队武装 DNA 鉴定中心规定不能根据单个碱基不同而得出排除结论。

5. mtDNA 序列单倍型数多，但是现在调查的人群数据少，还未获得真正有代表性的单倍型的群体数据，很难进行个体同一认定和亲子鉴定的概率计算。

6. mtDNA 序列分析方法复杂、繁琐。工作量大，费用昂贵，出错率高。Holland 和 Parsons 报道，mtDNA 测序错误率为 1% ~ 2%。

15.6.3 分型结果解释

进行物证样品比对时，首先要检查提取试剂对照或 PCR 阴性对照序列。如果提取试剂对照或 PCR 阴性对照得到与物证样品相同的序列，那么物证样品的结果绝对不能使用，必须重新分析。如果对照样品出现的序列与物证样品的不同，结果解释时必须结合所有实验结果，包括阳性对照、其他阴性对照、两次分析的结果和阴性对照相对于物证样品的相对强度。每次实验的阳性对照的结果必须是正确的，本次实验结果方可应用。

在确保 mtDNA 分型正确情况下，比较物证样品与参照样品序列。根据人群调查发现，无关个体在高变区Ⅰ和Ⅱ一般有4个以上碱基序列不同，异质性很少同时发生在两个碱基中，不是来自于同一母系的那些物证与比对样品间不止在一个核苷酸上有差异。

1. 如果比对样品有2个以上的碱基序列不同，可以排除他们的相关性。

2. 如果比对样品序列只有一个碱基差异。分两种情况：(1) 如果物证样品观察到异质体，而没有在已知样品中观察到，或反过来，在已知样品中观察到，而在物证样品中没有观察到，这种情况下不能排除他们有同一母系关系或来自于同一来源。(2) 如果未知样品和比对样品间只有一个碱基差异，而又没有观察到异质性的存在，此时的结果解释应是不能确定 (inconclusive)。已发现在个别同一母系人员间存在一些差异，有时不同母系或来源个体出现单核苷酸不同。

总之，样品间碱基差异数越大，他们来自于同一来源或有相同母系起源的可能性越小。

3. 如果序列相同，那么不能排除物证样品与参照样品来自于同一个来源的可能。比如现场提取的毛发与某一嫌疑人 A 的相同，那么这种匹配有三种可能情况：(1) 毛发可能来自于嫌疑人；(2) 毛发可能来自与嫌疑人有同一母系的亲属；(3) 毛发可能来自与嫌疑人无母系关系的无关个体。

如何判定结果属于哪种情况？属于其中第三种情况的可能性有多大？也就是说，具有这种毛发序列的潜在人群占多少？这一般取决于所分析的案件的案情以及所涉及到的 mtDNA 序列。如果案件限定在一定范围内，那么 mtDNA 序列相匹配可以认为是有决定意义的。如两个比对样品中同时观察到相同的异质体，这就增加了物证的证据力。又如，在飞机失事的空难中，采集的残骸是登机的乘客与机组成员的，也就是说，mtDNA 序列比对局限在一个比较固定有限的范围内。如果某一乘客的参考样品（乘客的母系家属）与他的序列相同，但与所有其他人员的序列不同，那么这种相匹配可以认为是同一的。

在实际工作中，更常见的是，进行 mtDNA 分析的案件物证的来源是未知的，没有任何关联，那么在这种情况下的 mtDNA 序列相同的资料是不能作为最终结论的决定性依据。

物证样品，要考虑到它的复杂性，如量少、存在降解或污染等各种因素，并且无法预先正确预测这些因素，因此在检验结果会出现原本来自于同一来源的样品的 DNA 图谱可能不同，一个图谱的带多，另一个的带少，但这种图谱的带不可能由一种型转变成另一种型的带，如等位基因 14，它不可能变成等位基因 10。造成这种同一来源样品的图谱有所差异的表现形式和产生原因在不同系统的遗传标记是不同的，如 RFLP 系统，可能出现：（1）带的漂移，量的差异或 DNA 中存在污染物如盐；（2）多带，内切酶的星号活性或污染有其他 DNA 或存在降解；（3）少带，存在 DNA 降解。PCR 系统可能出现：（1）少带，量少造成较小等位基因优先扩增，大等位基因丢失，或存在降解或抑制剂；（2）多带，量不平衡，扩增特异性降低，存在 Stutter 带，污染或多色荧光标记系统，不同染料间的干扰引起 pull - up 带。线粒体 DNA 系统的二异质性导致同一个体不同细胞间的差异，具体已在各标记有关部分详细阐述了。在排除时一定要注意这些问题。

经验上，一般以 9 个 STR 基因座检验可以达到个体同一认定，但必须注意已发现同胞之间在 Profilr plus 系统有 7 个基因座的分型完全一致，在其他两个基因座各有一个等位基因相同。因此当物证样品由于客观条件不能得到所有基因座的分型结果，在下结论时一定要慎重，虽图谱相同，但物证有可能来自于比对样品的同胞，而不是比对样品本人，表 20 - 2 是 13 个基因座在亲属间的匹配概率。

表 20 - 2　13 个基因座的在亲属间的匹配概率

人群	最常见杂合子				最常见纯合子			
	同胞		姨表姑舅叔侄		同胞		姨表姑舅叔侄	
	估计值	95%CI	估计值	95%CI	估计值	95%CI	估计值	95%CI
牙买加人	6.80×10^4	5.92×10^4	6.33×10^9	4.03×10^9	5.18×10^4	4.28×10^4	4.70×10^{10}	2.30×10^{10}
非裔美国人	8.24×10^4	7.13×10^4	1.23×10^{10}	7.59×10^9	5.62×10^4	4.60×10^4	6.38×10^{10}	3.46×10^{10}
欧裔美国人	7.78×10^4	6.76×10^4	1.16×10^{10}	7.23×10^9	6.93×10^4	5.69×10^4	1.39×10^{11}	7.51×10^{10}
瑞士	8.38×10^4	7.29×10^4	1.47×10^{10}	9.19×10^9	8.31×10^4	6.84×10^4	2.45×10^{11}	1.33×10^{11}
意大利	7.05×10^4	6.17×10^4	7.11×10^9	4.60×10^9	8.37×10^4	6.94×10^4	2.26×10^{11}	1.26×10^{11}
中国人	5.95×10^4	4.94×10^4	4.42×10^9	2.41×10^9	5.56×10^4	4.19×10^4	6.70×10^{10}	2.98×10^{10}
阿帕切人	2.63×10^4	2.29×10^4	4.84×10^8	3.14×10^8	8.43×10^3	7.00×10^3	3.11×10^8	1.84×10^8

注：匹配概率等于人群人数（N）分之一，N 为表中的数。

吕德坚等对 31 个 STR 基因座进行精液与血液分型比较，发现同一个体的精液与血液之间的基因型在 DXS6799、D8S1179 和 FGA 基因座不同（表 20 - 3），血液比精液的带多。因此个体同一认定时一定要注意，一个基因座的不同不能排除，否则会发生假排除。

表 20-3　三个不同个体的 3 个 STR 基因座的分型结果

个体	基因座	基因型	
		血液	精液
1	DXS6799	11，13	10
2	D8S1179	11，14，15	11，14
3	FGA	25.2，30.2	25.2

对于父系遗传的 Y 染色体和母系遗传线粒体 DNA 的遗传标记系统，由于它的遗传特点，不可能达到个人同一认定，同一父系来源的所有个体的 Y 染色体分型完全一致，同一母系的线粒体分型一致。对于此类标记，在排除比对样品的同一父系所有其他个体前提下，当物证样品与比对样品 Y 染色体标记的偶合率达到一定的个体认定水平时，可以结合案情如时间、地点、目击证人等认定他们来源于同一个体。同样地，在排除比对样品的同一母系所有其他个体前提下，物证样品与比对样品线粒体 DNA 标记的偶合率达到一定的个体认定水平，可以认定他们来源于同一个体。否则不能同一认定，不能肯定物证来源于比对样品还是来源与比对样品同一父系如兄弟、堂兄弟或同一母系的其他个体如姐妹、表兄妹等。线粒体 DNA 的异质性问题，同一个体的不同细胞也存在差异，一个碱基的差异不能排除他们来源于同一母系。分析时尽量采用同一种细胞，如血液与血液，毛发与毛发比对。

由于癌变，同一个体的正常组织与癌变组织的图谱不同，在 STR 系统有人已发现使等位基因重复序列减少的现象，即等位基因变小。应避免使用癌变组织器官进行 DNA 分型。

注意输血或骨髓或造血干系胞移植的个体，输血者要求在输血 9 周后进行 DNA 检验，骨髓或造血干系胞移植的个体的血液 DNA 分型已变为供体的分型，肌肉中等可能含有血管的组织 DNA，由于有血可能是供体和受体的混合分型，因此尽量采用同一种组织或细胞比对，对于有移植史的个体，最好采用其唾液进行 DNA 分型。

中国是一个多民族的国家，遗传标记的等位基因频率在各民族分布可能不同，但要注意：在释解结果时，嫌疑人的种族或民族是不相关的。目前也不可能通过法医 DNA 检验用的标记来确定一个人的种族或民族。选择物证图谱相关群体频率是受外部因素和案情如目击证人陈述或犯罪现场位置等影响的，当缺乏外部信息时，采用一般群体的频率来评价物证。

个体识别的物证经常为 2 个或多个来源 DNA 混合的混合样品，比如受害人的阴道上皮细胞和罪犯的精子混合物，因搏斗造成受害人和罪犯两人出血，在现场留下受害人的血液与罪犯血液混合，使 DNA 分型更加复杂，一定要予以注意。

混合样品有不同个体不同种类细胞混合的混合样品，如女性阴道分泌物和男性精液的混合斑，还有不同个体的同一类细胞如血液和血液混合的混合物。前者可以根据不同细胞对细胞裂解性不同应用差异裂解法分开，而后一种情况，得到的图谱为混合图谱，偶合率计算和结果解释不同于单一个体 DNA 图谱。如果为女性与男性 DNA 混合的混合物，可以

通过 Y 染色体的遗传标记进行分型，可以将男性 DNA 图谱变成为单一个体的图谱，结果变得简单些，如果男－男混合的样品，进行 Y 染色体的标记分析，结果也会变得简单些，男性个体在一基因座也只有一条带，对于女、女混合的混合物，目前尚无较为简单的方法。

对于已知是混合样品的物证，可以直接采用合适的分型系统进行分析，对于预先未知的物证，先采用常规方法，待结果发现为混合物时，再采用混合物适合的分型系统，以降低结果解释的复杂性，如果所有标记不能排除比对样品，需要进行偶合率（似然率）的计算。

20.4 混合样品分析

法医 DNA 分析的样品中有的可能是存在有来自多个供者的混合物，尤其是性犯罪案件中，一名或数名罪犯的 DNA 与受害人的 DNA 混合或数名罪犯的 DNA 混合。通常在一个或多个 STR 基因座出现 3 条或 3 条以上谱带，可以考虑是混合物（图 20－1）。当混合样品中的各个（2 个或 2 个以上）样品在特定的基因座上具有相同的等位基因，从谱带的数目上（即定性分析）无法确定是否为混合样品，用荧光方法定量检测时会发现这个峰高或面积有增大现象（图20－2），指示样品可能为混合样品。所以，定量的荧光检测有助于混合样品的检测分析。

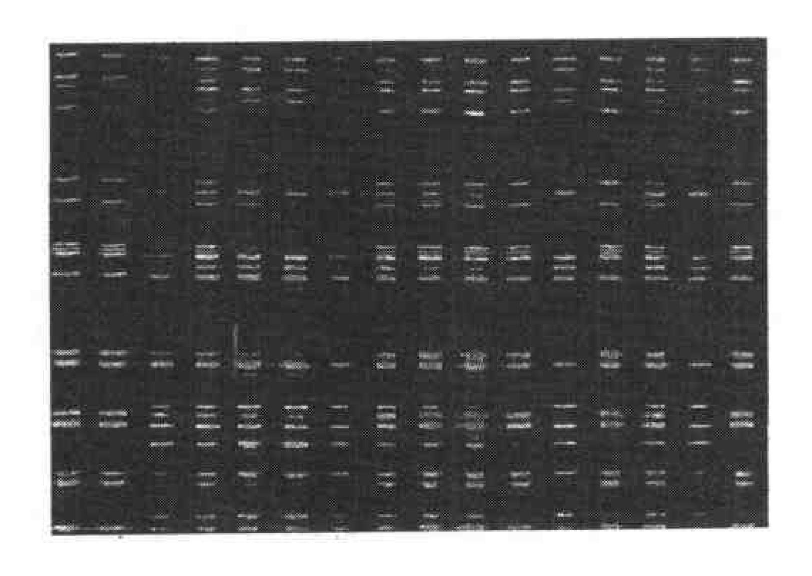


图 20－1 混合样品图谱

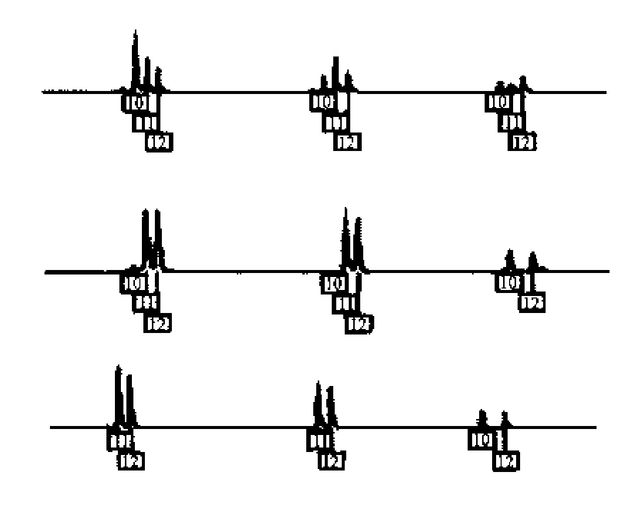


图 20－2　混合样品峰面积不平衡

a. 混合比例 1：1；b. 混合比例 10：1；c. 混合比例 20：1

1. 定性分析

二组分的混合样品在日常检验中较为常见。在分析二组分混合物时，在一个基因座上有四个等位基因，按相应组合，可有 3 对等位基因组合；同样地，在一个基因座有三个等位基因，可有 6 对组合；有 2 个等位基因，可有 4 种可能的组合；而一个等位基因，则只有一种可能，即二个个体均为纯合子且等位基因相同（表 20－4）。

表 20－4　二组分混合物在一个基因座图谱等位基因可能组合

等位基因数目及分型	组合
四个等位基因 A、B、C、D	A/B + C/D；A/C + B/D；A/D + B/C；
三个等位基因 A、B、C	A/A + B/C；B/B + A/C；C/C + A/B；
	A/B + A/C；A/B + B/C；A/C + B/C；
二个等位基因 A、B	A/B + A/B；A/A + B/B；
	A/B + A/A；A/B + B/B；
一个等位基因 A	A/A + A/A

2. 定量分析

荧光标记 PCR 产物可以准确测定各等位基因荧光峰高、峰面积，对各等位基因片段进行定量分析。任一基因座等位基因峰面积比值与混合样品内组分比例有密切的相关性。单一个体成分时，杂合子基因座有二个峰，即二个等位基因，峰面积基本接近；不同比例混合的混合样品会出现峰面积的不平衡。样品等位基因峰面积比值接近样品混合比例，可以由峰面积比值推知混合比例。有实验表明：当二个样品混合比例为 10:1 时，较少成分样品的等位基因峰高相当于主要等位基因的“stutter”峰的峰值；当混合比例为 20:1 时，基本上检测不到来自少量混合成分的等位基因，表现为单一组分图谱（图 20－2 b，c）。当确定为混合样品时，两两比较，计算两峰峰面积比，以确定各自可能的基因型。

3. 混合样品 STR 图谱的结果解释程序

实际混合样品结果分析程序可分为如下 5 个步骤：

(1) 确定是否为混合样品

根据得到的 STR 图谱中出现的峰或带的数目与峰面积的平衡情况确定、鉴定混合物的存在。

单一组分在一个基因座最多拥有二条带或二个峰，二条带或峰以外的均称为额外带或峰。在任何一个基因座出现三条或多条谱带通常直接表示为混合 STR 图谱。

一些情况下，尤其是那些相关个体是近亲的案件，或只得到部分图谱，或案件涉及的个体均为纯合子时，混合样品很可能没有额外带。尽管一个基因座没有额外带，但有时也可以从等位基因峰面积不对称现象推断出为混合 STR 图谱，如 Amelogenin XY 峰的严重不平衡，即表示可能为混合样品。

(2) 鉴定混合物中供体的数目

一旦确定为混合 STR 图谱，就要评估混合物中可能的供体数目。每一个基因座的等位基因数目及他们的相对比例是十分有用的参数。二个杂合子个体的最大等位基因数目为 4，在一个基因座出现 5 或 6 个等位基因则表明有三个或三个以上供体存在。根据经验，在办案中经常遇到的是二个个体混合的混合物，有时也会遇到三个个体混合的，应予注意。

(3) 确定混合样品中各混合组分的大约比例

对于荧光检测系统，计算各峰面积之比，确定混合组分的可能比例。

(4) 确定混合样品中各组分可能的等位基因对

综合整个 STR 图谱分析结果，确定混合组分比例，重新确定分出等位基因对。在实际应用中，有时分析人员已知其中一组分等位基因型是来自被害人，那么根据已知等位基因型在混合图谱中比对时将其减去，留下的则为另一组分的等位基因型。

(5) 比较混合样品可能组分组合图谱与来自比对样品的图谱

根据所推断的混合组分的等位基因型，与嫌疑人的等位基因型比对，如果在所有的基因座上都一致，那么此人就不可能被排除为混合斑痕一组分的可能供体，另一嫌疑人的等位基因型与混合样品中另一组分等位基因型（即剩下的等位基因型）匹配，则不能排除此二个嫌疑人为混合样品中 DNA 的供体，计算似然比率，确定现场样品是否来源于他们。

4. 似然比率（LR）的计算

混合样品 LR 的计算有以下 5 个步骤：

(1) 明确控方和辩护方的假设，即 Hp 和 Hd 含义；

(2) 在 Hp 下，计算 Pr（物证 | Hp)；

(3) 在 Hd 下，列出案犯可能的基因型与频率，计算 Pr（物证 | Hd)

(4) 计算 $LR = \frac{Pr（物证|Hp）}{Pr（物证|Hd）}$

(5) 计算多个独立分析基因座或标记的总 LR，

总 $LR = \prod_{i=1}^{n} LRi$

式中 LRi 表示某一个基因座的 LR，n 为总基因座或标记数。

根据案件性质、样品混合组分组成，二组分混合样品可分为三种情况：(1) 来自受害

人和嫌疑人的混合物；(2) 来自嫌疑人和另一未知个体的混合物；(3) 来自于二个嫌疑人的混合物。三种混合物的 LR 计算方法不同，下面将分别介绍。

(1) 来自于受害人和嫌疑人的混合物

Hp 表示混合斑为受害人和嫌疑人所留。

Hd 表示混合斑为受害人和人群中随机一个体所留。

$$LR = \frac{Pr（物证|Hp）}{Pr（物证|Hd）}$$

那么似然比率含义是混合斑为受害人和该嫌疑人所留的假设与混合斑为受害人和人群中随机个体所留的假设比值。

$$LR = \frac{P(M, V, S|Hp)}{P(M, V, S|Hd)} = \frac{P(M|V, S, Hp)}{P(M|V, S, Hd)} \times \frac{P(V, S|Hp)}{P(V, S|Hd)} = \frac{P(M|V, S, Hd)}{P(M|V, S, Hd)}$$

式中 M、V、S 分别表示混合物、受害人和嫌疑人基因型

①混合 DNA 样品为四等位基因 A1、A2、A3、A4，受害人基因型为 A3、A4，嫌疑人为 A1、A2（如图 20－3)。

	M	V	S
A_1	—		—
A_2	—		—
A_3	—	—	
A_4	—	—	

图 20－3　混合物受害人和嫌疑人的基因型

在 Hp 下，P（M|V, S, Hp）＝1；

在 Hd 下，案犯 S 基因型（G_S）为 A_1A_2，频率为 $2P_1P_2$，P（M|V, S, Hd）＝$2P_1P_2$；

$$LR = \frac{P(M|V, S, Hp)}{P(M|V, S, Hd)} = \frac{1}{2P_1P_2}$$

例如　混合物 M 的基因型为 G_m = 14, 15, 17, 18，受害人 G_v = 15, 18，嫌疑人 G_s = 14, 17，p_{14} = 0.033，p_{17} = 0.239。

$$LR = \frac{1}{2P_{14}P_{17}} = \frac{1}{2 \times 0.033 \times 0.239} = 63.4$$

②三等位基因（A_1、A_2、A_3）

三等位基因有以下二种情况：

1) 受害人为纯合子，即 $G_M = A_1、A_2、A_3$，$G_v = A_3$，$G_s = A_1A_2$（图 20－4a）

	(a) M	(a) V	(a) S	(b) M	(b) V	(b) S
A_1	—		—	—		—
A_2	—		—	—	—	
A_3	—	—		—	—	

图 20－4　混合物受害人和嫌疑人的基因型

在 Hp 下，P（M|V, S, Hp）＝1

在 Hd 下，$G_S = A_1A_2$

$LR = 1/(2P_1P_2)$

2）受害人为杂合子，即 $G_v = A_2A_3$，$G_s = A_1$（图 20－4b）

在 Hp 下，P（M|V，S，Hp）＝1

在 Hd 下，案犯基因型可能是 A_1A_2，A_1A_3，A_1A_1

P（M|V，S，Hd）$= 2P_1P_2 + 2P_1P_3 + P_1^2$

$$LR = \frac{1}{2P_1P_2 + 2P_1P_3 + P_1^2}$$

③混合物为二等位基因 A1、A2

1）受害人为杂合子（图 20－5a）

	M	V	S	M	V	S	
A_1	—	—	—	—		—	—
A_2	—	—	—	—	—	—	A_2
	a				b		

图 20－5 混合物受害人和嫌疑人的基因型

在 Hp 下，P（M|V，S，Hp）＝1；

在 Hd 下，案犯 S 为基因型（G_s）为 A_1A_2、A_1A_1、A_2A_2，频率为 $2P_1P^2$、P_1^2、P_2^2，P（M|V，S，Hd）$= 2P_1P_2 + P_1^2 + P_2^2$

$$LR = \frac{P(M|V, S, Hp)}{P(M|V, S, Hd)} = \frac{1}{2P_1P_2 + P_1^2 + P_2^2};$$

2）受害人为纯合子，基因型为 A_2A_2（图 20－5b）

在 Hp 下，P（M|V，S，Hp）＝1；

在 Hd 下，案犯 S 为基因型（G_S）为 A_1A_2、A_1A_1，频率为 $2P_1P_2$、P_1^2，

P（M|V，S，Hd）$= 2P_1P_2 + P_1^2$；

$$LR = \frac{P(M|V, S, Hp)}{P(M|V, S, Hd)} = \frac{1}{2P_1P_2 + P_1^2}$$

受害人和嫌疑人组成的混合物 LR 计算总结如表 20－5 所示。

表 20－5 来自于受害人和嫌疑人混合物

混合斑基因型	受害人	嫌疑人	案犯可能基因型	P（分子）	P（M|V，S，Hd）	LR
$A_1A_2A_3A_4$	A_3A_4	A_1A_2	A_1A_2	1	$2P_1P_2$	$1/(2P_1P_2)$
$A_1A_2A_3$	A_2A_3	A_1	A_1A_2，A_1A_3，A_1A_1	1	$2P_1P_2 + 2P_1P_3 + P_1P^2$	$1/(2P_1P_2 + 2P_1P_3 + P_1P^2)$
	A_3	A_1A_2	A_1A_2	1	$2P_1P_2$	$1/(2P_1P_2)$
A_1A_2	A_1A_2	A_1A_2	A_1A_2，A_1A_1，A_2A_2	1	$2P_1P_2 + P_1P^2 + P_2^2$	$1/(P_1 + P_2)^2$
	A_1A_2	A_1	A_1A_2，A_1A_1，A_2A_2	1	$2P_1P_2 + P_1^2 + P_2^2$	$1/(P_1 + P_2)^2$
	A_1	A_1A_2	A_1A_2，A_2A_2	1	$2P_1P_2 + P_2^2$	$1/(2P_1P_2 + P_2^2)$
	A_1	A_2	A_1A_2，A_2A_2	1	$2P_1P_2 + P_2^2$	$1/(2P_1P_2 + P_2^2)$

(2) 嫌疑人和未知个体混合

Hp 表示混合斑来自于嫌疑人和另一未知人 C_1；

Hd 表示混合斑来自于二个随机个体 C_1 和 C_2。

$LR = P_r(M|S, Hp) / P_r(M|S, Hd)$

$P_r(M|S, Hd) = \sum P_r(M|S, C_i, H_p)\ P_r(C_i|S, Hp)$

$P_r(M|S, Hd) = \sum_{i}^{i=1}\sum_{j} P_r(M|C_i, C_j, Hd), P_r(C_i, C_j|Hd) = \sum_{i}^{i=1}\sum_{j} P_r(M|C_i, C_j, Hd)$

(因为每一 C_i 与 C_j 组合的 $P_r(M|C_i, C_j, Hd) = 1$)

$$LR = \frac{\sum P_r(C_i|S, Hp)}{\sum_i\sum_j P_r(C_i, C_j|Hd)}$$

①混合物为四等位基因 A_1、A_2、A_3、A_4，$G_q = A_1A_2$（图 20－6）

在 Hp 下，S 为 A_1A_2，则 C 为 A_3A_4，$P_r(M|S, Hp) = 1 \times 2P_3P_4 = 2P_3P_4$；

在 Hd 下，C_1 和 C_2 均有以下 6 种组合基因型（表 20－6）。

	M	S
A_1	—	—
A_2	—	—
A_3	—	
A_4	—	

图 20－6　混合物、嫌疑人基因型

表 20－6　C_1 和 C_2 可能基因型与 $P_r(C_i, C_j|Hd)$

C_1	C_2	$P_r(C_i, C_j\|Hd)$
A_1A_2	A_3A_4	$2P_1P_2 \times 2P_3P_4$
A_1A_3	A_2A_4	$2P_1P_3 \times 2P_2P_4$
A_1A_4	A_2A_3	$2P_1P_4 \times 2P_2P_3$
A_2A_3	A_1A_4	$2P_2P_3 \times 2P_1P_4$
A_2A_4	A_1A_3	$2P_2P_4 \times 2P_1P_3$
A_3A_4	A_1A_2	$2P_3P_4 \times 2P_1P_2$

$P_r(M|S, Hd) = \sum_i\sum_j P_r(M|C_i, C_j, Hd) = 24\ P_1P_2\ P_3P_4$

$$LR = \frac{2P_3P_4}{24\ P_1P_2\ P_3P_4} = \frac{1}{12\ P_1P_2}$$

②混合物为三等位基因 A_1、A_2 和 A_3

三等位基因混合物有二种情况：

1）嫌疑人为杂合子 A_1A_2（如图 20－7）

M	S	C		
—	—	—		
—	—		—	
—		—	—	—

图 20－7 混合物、嫌疑人基因型

在 Hp 下，C 有三种可能基因型 A_1A_3、A_2A_3、A_3A_3

$\Pr(M|S,\ Hp) = 2P_1P_3 + 2P_2P_3 + P_3^2$

在 Hd 下，二随机个体有基因型及概率如表 20－7。

表 20－7 C_1 和 C_2 可能基因型、概率

C_i	C_j	P_r（Ci，Cj|Hd）	C_i	C_j	P_r（C_i，C_j|Hd）
A_1A_2	A_1A_3	$2P_1P_2 \times 2P_1P_3$	A_2A_3	A_1A_2	$2P_2P_3 \times 2P_1P_2$
	A_2A_3	$2P_1P_2 \times 2P_2P_3$	A_2A_3	A_1A_3	$2P_2P_3 \times 2P_1P_3$
	A_3A_3	$2P_1P_2 \times P_3^2$		A_1A_1	$2P_2P_3 \times P_1^2$
A_1A_3	A_1A_2	$2P_1P_3 \times 2P_1P_2$	A_1A_1	A_2A_3	$P_1^2 \times 2P_2P_3$
	A_2A_3	$2P_1P_3 \times 2P_2P_3$	A_2A_2	A_1A_3	$P_2^2 \times 2P_1P_3$
	A_2A_2	$2P_1P_3 \times 2P_2^2$	A_3A_3	A_1A_2	$P_3^2 \times 2P_1P_3$

$P_r(M|S,\ Hd) = \sum_i\sum_j P_r P_r(M|C_i,\ C_j,\ Hd) = 12P_1P_2P_3(P_1+P_2+P_3)$

$$LR = \frac{2P_1P_3 + 2P_2P_3 + P_3^2}{12P_1P_2P_3(P_1+P_2+P_3)} = \frac{2P_1 + 2P_2 + P_3}{12P_1P_2(P_1+P_2+P_3)}$$

2）嫌疑人为纯合子 A_1A_1（图 20－8）

	M	S	C
A_1	—	—	
A_2	—		—
A_3	—		—

图 20－8 混合物、嫌疑人基因型

在 Hp 下，C 有一种基因型 A_2A_3，$P_r(M|S,\ Hd) = 2P_2P_3$

在 Hd 下，二随机个体基因型及概率同嫌疑人杂合子的情况

$\Pr(M|S,\ Hd) = 12P_1P_2P_3(P_1+P_2+P_3)$

$$LR = \frac{2P_2P_3}{12P_1P_2P_3(P_1+P_2+P_3)} = \frac{1}{6P_1(P_1+P_2+P_3)}$$

③混合斑为二等位基因

1）嫌疑人为杂合子 A_1A_2（图 20-9）

在 Hp 下，C 基因型可能为 A_1A_2，A_1A_1 和 A_2A_2

$P_r(M|S, Hp) = 2P_1P_2 + P_1^2 + P_2^2$

在 Hd 下，二随机个体基因型、概率如表 20-8。

表 20-8　C_1 和 C_2 可能基因型 $P_r(C_i, C_j|Hd)$

C_1	C_2	$P_r(C_i, C_j\|Hd)$
A_1A_2	A_1A_1	$2P_1P_2 \times P_1^2$
	A_1A_2	$2P_1P_2 \times 2P_1P_2$
	A_2A_2	$2P_1P_2 \times P_2^2$
A_1A_1	A_1A_2	$P_1^2 \times 2P_1P_2$
	A_2A_2	$P_1^2 \times P_2^2$
A_2A_2	A_1A_2	$P_2^2 \times 2P_1P_2$
	A_1A_1	$P_2^2 \times P_1^2$

$$P_r(M|S, Hp) = 2P_1^3P_2 + 4P_1^2P_2^2 + 2P_1P_2^3 + 2P_1^3P_2 + P_1^2P_2^2 + 2P_1P_2^3 + P_1^2P_2^2$$
$$= 2P_1P_2(3P_1P_2 + 2P_1^2 + 2P_2^2)$$

$$LR = \frac{2P_1P_2 + P_1^2 + P_2^2}{2P_1P_2(3P_1P_2 + 2P_1^2 + 2P_2^2)}$$

	M	S	C	
A_1	—	—	—	
A_2	—	—	—	—

图 20-9　混合物嫌疑人和未知个体的基因型

2）嫌疑人为纯合子 A_1A_1

在 Hp 下，S 为 A_1A_1，C 基因型可能为 A_1A_2，A_2A_2，$P_r(M|S, Hp) = 2P_1P_2 + P_2^2$

在 Hd 下，二随机个体基因型、概率及同杂合子个体，

$P_r(M|S, Hp) = 2P_1P_2(3P_1P_2 + 2P_1^2 + 2P_2^2)$

$$LR = \frac{2P_1 + P_2}{2P_1(3P_1P_2 + 2P_1^2 + 2P_2^2)}$$

表 20-9 为来自于嫌疑人和另一未知个体的混合样品的 LR 计算简表。

表 20-9 混合样品来自于嫌疑人和另一未知个体的 LR

混合斑基因型	嫌疑人基因型	未知个体基因型	LR
$A_1A_2A_1A_2$	A_1A_2	A_1A_2	$1/(2P_1P_2)$
$A_1A_2A_3$	A_1A_2	A_1A_2，A_2A_3，A_3A_3	$(2P_1+2P_2+P_3)/[12P_1P_2(P_1+2P_2+P_3)]$
	A_1A_1	A_2A_3	$1/[6P_1(P_1+P_2+P_3)]$
A_1A_2	A_1A_2	A_1A_1，A_2A_3，A_3A_3	$(2P_1P_2+P_1^2+P_2^2)/[2P_1P_2(3P_1P_2+2P_1^2+2P_2^2)]$
	A_1A_1	A_1A_2，A_2A_2	$(2P_1+P_2)/[2P_1(3P_1P_2+2P_1^2+2P_2^2)]$

（3）混合斑为二个嫌疑人

Hp 表示混合斑为来自于二嫌疑人 S_1 和 S_2。不论何种基因型，$P_r(M|S_1S_2, Hp)=1$；

Hd 表示混合斑来自于未知的二随机个体 C_1 和 C_2。$P_r(M|S_1S_2, Hp)$ 与来自于嫌疑人与未知个体的混合物包含相应等位基因的 $Pr(M|S_1S_2, Hp)$ 相同。

①四等位基因 $M=A_1$、A_2、A_3、A_4，嫌疑人 $S_1=A_1A_2$，$S_2=A_3A_4$

在 Hp 下，$P_r(M|S_1S_2, Hp)=1$

在 Hd 下，$P_r(M|S_1S_2, Hd)$ 同来自已知一嫌疑人和未知个体的四等位基因混合物，即 $Pr(M|S_1S_2, Hd)=24P_1P_2P_3P_4$

$$LR=\frac{1}{24P_1P_2P_3P_4}$$

②三等位基因

不论二嫌疑人基因型是纯合子还是杂合子（图 20-10a，b），LR 计算相同

a

	M	S_1	S_2
A_1	—	—	
A_2	—		—
A_3	—		—

b

	M	S_1	S_2
A_1	—	—	
A_2	—	—	—
A_3	—		—

图 20-10 混合物、二嫌疑人基因型

在 Hp 下，$P_r(M|S_1S_2, Hp)=1$

在 Hd 下，二未知随机个体基因型，概率与来自一嫌疑人与未知随机个体混合物情况相同，$P_r(M|S_1S_2, Hd)=12P_1P_2P_3(P_1+P_2+P_3)$

$$LR=\frac{1}{12P_1P_2P3(P_1+P_2+P_3)}$$

③二等位基因

二等位基因混合物基因型如图 20-11a，b，c 所示，三种情况的 LR 计算相同。

	M	S_1	S_2		M	S_1	S_2	S_1	S_2
A_1	—	—	—	A_1	—		—	—	
A_2	—	—	—	A_2	—	—	—		—
		a				b		c	

图 20-11　混合物、二嫌疑人基因型

在 Hp 下，三种情况 P_r（$M|S_1S_2$，Hp）$=1$

在 Hd 下，二个未知随机个体基因型、概率 P_r（$M|S_1S_2$，Hd）与来自于一嫌疑人与另一未知个体和混合物的相同。

P_r（$M|S_1S_2$，Hd）$=2P_1P_2(3P_1P_2+2P_1^2+2P_2^2)$

$$LR=\frac{1}{2P_1P_2(3P_1P_2+2P_1^2+2P_2^2)}$$

表 20-10 为混合样品来自于 2 个嫌疑人的 LR 计算。

表 20-10　混合样品来自于二个嫌疑人的 LR

混合斑基因型	LR
A1、A2、A3、A4	$1/(24P_1P_2P_3P_4)$
A1、A2、A3	$1/[12P_1P_2P_3(P_1+P_2+P_3)]$
A1、A2	$1/[2P_1P_2(3P_1P_2+2P_1^2+2P_2^2)]$

(4) n 个独立遗传的基因座或标记的总 LR

n 个独立遗传的基因座或标记的 LR 值的乘积即系统的总 LR

$$总\ LR=\prod_{i=1}^{n}LR_i$$

i 代表分析的某一个基因座，n 为系统分析的基因座或遗传标记数。

第二十一章 亲子鉴定

21.1 概述

21.1.1 亲子鉴定案件

亲子鉴定是法医 DNA 分析应用的另一个领域。DNA 是遗传物质，稳定地从亲代传递给子代，通过 DNA 检验确定假设的父母与子女之间是否存在亲生血缘关系或假设个体与待检验的个体或物证之间是否存在亲缘关系，可以进行亲子鉴定或亲权鉴定。随着技术与经济发展，越来越多的案件涉及到亲子鉴定，亲子鉴定的案件大致有：

(1) 丈夫怀疑孩子非亲生。

(2) 非婚生育，男方怀疑孩子非亲生或女方想搞清孩子的生父究竟是谁。

(3) 强奸致孕，女方指控某男为孩子生父等。

(4) 认亲、移民、继承财产、上户口。

(5) 怀疑产房或育婴室调错婴儿。

(6) 计划生育确认孩子是否为该对夫妇所超生或领养。

(7) 被拐骗儿童或失散家庭成员的确认。

(8) 尸体（碎尸块、尸骨等）身份（尸源）确定。

(9) 需确定现场物证或犯罪嫌疑人衣服或家里斑迹是否为受害人所留，但受害人 DNA 已被销毁，无法得到；或如进行强奸案件中精斑同一认定，但嫌疑人在逃，通过亲子鉴定确认精斑是否为嫌疑人所留。

上述（1）～（3）中情况是标准的三联体亲子鉴定：参加的个体为母、子与假设父，母子关系已经确定无疑，要求鉴定假设父与孩子之间的亲子关系；其余 6 种情况涉及可疑父亲、母亲双方的确定。

母－子（儿子或女儿，下同）－男三方均参加检验的情况称为三联体（triplet），这样的鉴定为双亲检验。真实存在的母－子－父关系的三联体，称为真三联体，不存在母－子－父关系的三联体，称之为假三联体。其中被检验的男子称为假设父亲或被告父亲（alleged father）。如果父母均需要确定的，则称被检验的某男、某女为假设父亲和假设母亲，亦属于双亲检验。当受检者仅为假设父或假设母和孩子二方时，称之为二联体（biplet），这种只有父母一方的二联体的亲子鉴定称为单亲检验。反转亲子鉴定是指父母已确定，检验孩子是否为他们亲生的亲子鉴定，主要见于失踪者（活体、尸体）的身份确定。

亲子鉴定最早的应用对象是有争议的父母与子女之间的血缘关系的确定，随着越来越多的遗传标记系统的开发和应用，被检测的对象已不再局限于父母与子女两代之间的个体，对同胞、隔代、甚至旁系个体间是否存在亲缘关系也有可能进行相应的鉴定。如某种原因（如死亡），被假设的父亲不能参加检验，而由假设父的亲属参加检验。或者父或/和母已亡，或因某种原因不能参加检验，要求进行隔代、同胞或叔侄、姨甥等旁系人员间的

亲权鉴定。选择合适的标记可以进行上述鉴定。

21.1.2 **亲子鉴定应用的遗传标记**

应用于亲子鉴定的 DNA 遗传标记系统很多，按遗传方式可归纳分为 4 类：

1. 常染色体上基因座或遗传标记，按孟德尔遗传规律传递；

2. 线粒体 DNA 非编码区的多态性。按母系遗传方式传递，可以确定检验样品是否来自于同一母系，适用于那些父亲不能参加鉴定的母子间的单亲鉴定或同胞之间的或隔代或旁系亲缘关系鉴定，如再结合按孟德尔遗传的遗传标记检验，可以正确确定样品间的亲缘关系；

3. Y－染色体上基因座或遗传标记。按父系遗传方式传递，可以确定检验样品是否来自于同一父系，适用于那些母亲不能参加鉴定的父子间的单亲鉴定或男性同胞之间或隔代或旁系的亲缘关系鉴定，如再结合按孟德尔遗传的遗传标记检验，可以正确确定样品间的亲缘关系；

4. X－染色体上基因座或遗传标记。由于遗传的定向性，它适合三联体的亲子鉴定或除父子关系外的其他单亲的亲缘鉴定。

21.1.3 **亲子鉴定原理**

据孟德尔遗传的分离和自由组合定律，亲代基因型决定子代基因型，他们之间基因型关系如表 21－1。在没有基因突变、分型错误的前提下：①孩子的一对等位基因必定是一个来自父亲，一个来自母亲；②孩子不可能带有双亲均没有的等位基因。这两点是亲子鉴定的基本原理。也就是说，在肯定孩子的某个等位基因为生父基因，而假设父并不带此等位基因时，可排除他为孩子的生父；在肯定孩子的某个等位基因来自于其父亲，而假设父也带有此等位基因时，不能排除他为孩子的生父。对于父系遗传的 Y 染色体的标记，子代的分型必定与父亲的相同，而且同一父系的所有个体的分型一致。母系遗传的线粒体 DNA，子代的分型与母亲的分型一致，并且同一母系的所有个体的分型一致。仅根据上述遗传定律假设父与孩子之间在一个遗传标记上不符合遗传规律可以排除其父子关系，但由于突变的存在，一个基因座不符合遗传规律，可能是由突变造成的，而且有文献已报道在真实的家系中存在两个遗传标记的突变，因此一个遗传标记的不符不能否定其亲子关系，要求在 2 个以上不同基因座同时不符才能否定其亲子关系。

表 21－1　亲代与子代基因型关系

亲代基因型组合	子代基因型（机会）
aa × aa	aa（1）
aa × bb	ab（1）
ab × cd	ac（1/4），ad（1/4），bc（1/4），bd（1/4）
aa × bc	ab（1/2），ac（1/2）
ab × ab	aa（1/4），bb（1/4），ab（1/2）
aa × ab	aa（1/2），ab（1/2）
ab × ac	aa（1/4），ac（1/4），bc（1/4），ab（1/4）

* a、b、c、d 代表某一基因座位上不同的等位基因

21.2 非父排除概率与亲权指数

21.2.1 非父排除概率

各种遗传标记的多态性程度不同，鉴别能力有所差异，可以用非父排除概率（probability of excluding paternity，PEP）或排除概率（exclusion probability，EP）表示其在亲子鉴定中的鉴定能力。非父排除概率又称父权排除概率（excluding probability of paternity，EPP）指的是，通过检测某一个遗传标记系统，能将不是生父的假设父亲排除的机会，它是衡量一个遗传标记排除非父能力的一个客观标准，表示在所有非父被控为生父的男子中，用该标记否定父权有多大的可能性。非父排除概率大小与该标记的遗传方法、等位基因数及其在人群中的等位基因频率有关。不同遗传方式，EP 计算不同。

1. 离散型等位基因系统：如小卫星 VNTR 基因座、微卫星 VNTR 基因座、单基因座 VNTR 纹印等

（1）双亲检验母子关系已确定时，排除概率公式为：

$$EP=\sum_{i=1}^{n}p_i(1-p_i)^2(1-p_i+p_i^2)+\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}p_ip_j(p_i+p_j)(1-p_i-p_j)^2$$

或

$$EP=\sum_{i=1}^{n}p_i(1-p_i)^2-\frac{1}{2}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}(p_ip_j)^2(4-3p_i-3p_j)$$

式中 n 为等位基因数，p_i、p_j 分别为第 i、第 j 个基因频率。

近似计算公式：$EP=h^2(1-2hH^2)$

式中：H 为纯合子频率，h 为杂合子频率，$H+h=1$

（2）单亲鉴定非父排除概率计算公式为：

$$EP=\sum_{i=1}^{n}p_i^2(1-p_i)^2+\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}2p_ip_j(1-p_i-p_j)^2$$

2. Y 染色体遗传标记系统：

$$EP=\sum_{i=1}^{n}p_i(1-p_i)=1-\sum_{i=1}^{n-1}p_i^2$$

式中 n 为基因座的等位基因数，p_i 为各等位基因频率。

3. X 染色体遗传标记系统：由于 X 染色体的遗传特点，女孩可以否定父权或母权，男孩只能否定母权。

（1）母女关系肯定时，非父排除概率为：

$$EP=1-\sum_{i=1}^{n}p_i^2+\sum_{i=1}^{n}p_i^4-(\sum_{i=1}^{n}p_i^2)^2$$

或

$$EP=1-\sum_{i=1}^{n}p_i^2(1-p_i)+2\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}p_ip_j(1-p_i-p_j)$$

（2）单亲鉴定时，

①女孩与非父的排除概率为：

$$EP=\sum_{i=1}^{n}p_i(1-p_i)(1-p_i+p_i^2)+\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}p_ip_j(p_i+p_j)(1-p_i-p_j)$$

或

$$EP = 1 - 2\sum_{i=1}^{n} p_i^2 + \sum_{i=1}^{n} p_i^3$$

②女孩与非母的排除概率计算 同染色体基因组系统的单亲检验。

③男孩与非母的排除概率为

$$EP = \sum_{i=1}^{n} p_i \ (1 - p_i)^2$$

4. 累积非父排除概率

亲子鉴定一般使用多个遗传标记系统，因此求得每个系统的排除概率后尚需求得累积非父排除概率。多个互相独立的 DNA 标记系统的累积非父排除概率（cumulative chance of exclusion，CCE）为：

$$CCE = 1 - \prod_{i=1}^{n} \ (1 - EPi)$$

式中 n 为遗传标记数，EPi 为第 i 系统的非父排除概率

从上式中看出，检测遗传标记越多，累积排除概率越高，但当 CCE 接近 1 时，增加检测项目，CCE 值增加不明显，即 CCE 的增加渐趋平稳。

对于连锁系统的遗传标记，它们的累积非父排除率不是各个遗传标记的乘积，而是按其组成的单倍型频率计算非父排除概率。

非父排除概率反映了某个 DNA 标记系统在亲子鉴定中排除父权能力的大小，是选择亲子鉴定遗传标记的依据和衡量从事亲子鉴定实验室的质量控制标准之一。表 21 - 2 为目前常用 13 个 STR 基因座的非父排除率。累积非父排除概率越高，排除非父的机会亦越高，若假设父不被排除，肯定其为生父的机会也越高。为了最大限度排除非父，提高排除假设父的效率，选择排除概率高的遗传标记系统和选用更多的遗传标记是十分必要的。

表 21 - 2 中国汉族群体 13 个 STR 基因座的非父排除概率及累积非父排除概率

基因座	非父排除概率	累积非父排除概率
TPOX	0.302	0.302
D3S1358	0.510	0.65798
FGA	0.635	0.8751627
D5S818	0.683	0.9604266
CSF1PO	0.413	0.9767705
D7S820	0.445	0.9871077
D8S1179	0.685	0.995939
TH01	0.326	0.9972629
VWA	0.540	0.9987684
D13S317	0.593	0.9994988
D16S539	0.409	0.9997489
D18S51	0.667	0.9999164
D21S11	0.667	0.9999722

21.2.2 父权指数

遗传标记不能排除鉴定样品间的亲缘关系时，需要评价假设父与孩子有亲生关系可能

性大小（即亲权机会），以判定是否亲生关系。常用评价亲权机会的指标有父权指数(PI)、似然比率（LR）和父权相对机会等。

父权指数（paternity index，PI）又称亲子关系指数，是指假设父提供生父基因的可能性（X）与随机男子提供生父基因的可能性（Y）的比值，表示假设父为孩子生父比随机男子为孩子生父的可能性大多少倍，即 PI = X/Y。

假设父为生父的机会：

X = ∑母提供生母基因概率（f）×假设父提供生父基因概率（c）= ∑f×c

随机男子为生父的机会：

Y = ∑母提供生母基因概率（f）×随机男子提供生父基因概率（p）= ∑f×p

因此

$$PI = \frac{X}{Y} = \frac{\sum f \times c}{\sum f \times p} \qquad p \text{ 为人群中该基因频率}$$

PI 值越大，表示真三联的可能性大，PI 值越小，表示假三联的可能性大。

k 个独立遗传标记联合的 PI 为各个标记的乘积 PI，即

$$PI = \prod_{j=1}^{k} PI_j$$

k 为测定系统数，PIj 为各个系统的 PI 值。

21.2.3 似然比率

似然比率（likelihood ratio，LR）是另一个用于评价亲子关系的名词，以分型结果不违反遗传规律作为条件，假设父是孩子生父的事件概率与一个随机男人是孩子生父的事件概率之比为似然比。父权指数与似然比是同一概念，只是表达方式上有所区别，他们包含了确定亲子关系的所有信息，是一项重要的亲子关系参数。

Hp：假定父为孩子的生父

Hd：人群中随机男子为孩子的生父

则

$$LR = \frac{P(E \mid Hp)}{P(E \mid Hd)} = \frac{P(G_C, G_M, G_{AF} \mid Hp, I)}{P(G_C, G_M, G_{AF} \mid Hd, I)} = \frac{P(A_M \mid G_M)P(A_P \mid G_{AF}, Hp)}{P(A_M \mid G_M, Hd)P(Ap \mid G_{AF}, Hd)}$$

$$= \frac{P(AF)P(M)P(C \mid M, AF = F)}{P(AF)P(M)P(C \mid M)} = \frac{x}{y} = \text{PI}$$

式中 G_C，G_M，G_{AF}分别代表孩子、母亲和假设父的基因型，I 可以省略。

21.2.4 父权概率

父权概率（probability of paternity）又称父权肯定机率或亲子关系概率，以 W 表示。

以 X 表示给定群体中母－子－假设父真三联的机会，Y 表示假三联的机会。根据非遗传学证据估计的假设父是孩子生父的概率称为前概率（prior probability of paternity），以 P_0 表示，则 $1 - P_0$ 为非遗传学证据估计的假设父不是孩子生父的概率。

$$W = \frac{P_0 X}{P_0 X + (1 - P_0)\ Y} = \frac{(PI \times P_0)}{PI \times P_0 + (1 - P_0)}$$

当 $P_0 = 0.5$ 时，即非遗传学证据估计的假设父是孩子生父或不是生父的概率相同，则

$$W = \frac{X}{X+Y} = \frac{PI}{PI+1}$$

为了避免人为的估计因素影响亲子鉴定，一般将 P_0 定为 0.5。

21.2.5 PI **计算方法**

①根据母－子组合，确定来自母亲的基因（生母基因）和必定来自父亲的基因（生父基因）。

②计算母亲提供生母基因概率（f）和假设父提供生父基因概率（c）。

③计算假设父成为生父的机会：$x = \sum f \times c$

④计算随机男子成为生父的机会：$y = \sum f \times q$。随机男子提供生父基因概率（q）为人群中该基因频率。

⑤计算 PI

$$PI = \frac{X}{Y}$$

亲子鉴定类型或遗传标记系统不同，PI 或 LR 计算方法有所不同。

1. 三联体亲子鉴定

（1）母亲肯定情况下，经典三联体

母亲为杂合子 AB 时，母亲有二个等位基因 A、B 可提供，母亲传递生母基因的概率为 0.5，纯合子时只有一种等位基因可提供，母亲传递生母基因的概率为 1，同样，假设父为杂合子时假设父传递生父概率为 0.5，纯合子传递生父概率为 1。随机男子传递概率为该生父基因在人群中的频率。表 21－3 为 PI 简化计算表。

表 21－3　PI 简化计算

基因型组合			母亲确定情况			双亲皆疑		
子	母	假定父 F	X	Y	PI	X	Y	PI
AA	AA	AA	1	p	1/p	1	p^2	$1/p^2$
		AB	1/2	p	1/2p	0.5	p^2	$1/(2p^2)$
		BC	0	p	0	0	p^2	0
	AB	AA	1	p	1/p	0.5	p^2	$1/(2p^2)$
		AB	1/2	p	1/2p	0.25	p^2	$1/(4p^2)$
		AC	1/2	p	1/(2p)	0.25	p^2	$1/4p^2$
AB	AA	BB	1	p	1/p	1	2pq	1/2pq
		BC	1/2	p	1/2p	1/2	2pq	1/(4pq)
		AB	1/2	p	1/(2p)	1/2	p^2	$1/(2p^2)$
	AB	AA 或 BB	1/2	(p+q)/2	1/(p+q)	1/2	2pq	1/(4pq)
		AB	1/2	(p+q)/2	1/(p+q)	1/4	2pq	1/(4pq)
		BC	1/4	(p+q)/2	1/2(p+q)	1/4	2pq	1/(8pq)
	AC	BB	1	p	1/p	1/2	2pq	1/(8pq)
		BD	1/2	p	1/2p	1/4	2pq	1/(8pq)
		AB	1/2	p	1/(2p)	1/4	2pq	1/(8pq)

p、q 为人群生父基因频率

对于 Y 染色体遗传标记系统　决定男性性别的 Y 染色体系父系遗传，无论在父子亲权认定，还是在祖孙、叔侄亲权认定中，式中 x 均为 1，y 为随机人群该基因（如检测 1 个 Y－STR 基因座）或单倍型频率（如检测 2 个以上 Y－STR 基因座）。

PI＝1/y。

2．父母均怀疑的亲子鉴定

对于此类亲子鉴定 PI 计算目前还没有一个十分理想的算法。有些人认为按母亲肯定仅怀疑父亲的亲子鉴定计算；也有人将它按二个单亲计算，如果它们的 PI 值都大于 2 000，可以认定他们的亲权关系；还有些人认为 PI 计算步骤及原则同母亲肯定仅怀疑父亲的亲子鉴定，但式中分子为假设父遗传生父基因和假设母遗传生母基因产生孩子的机会，分母为随机父和随机母产生孩子的机会即为人群中孩子表现型频率。在极大多数情况下，这种算法得到的 PI 值较仅怀疑父亲的 PI 值大，有可能估计过高。

此种计算方法 PI 计算可归纳简化为：

（1）假设父和假设母均为纯合子时 x＝1；

（2）假设父和假设母一个为纯合子，一个为杂合子时 x＝0.5；

（3）假设父和假设母均为杂合子时 x＝0.25，但假设父、假设母、子为相同基因型杂合子时 x＝0.5。

（4）y 为人群中孩子基因型频率，子为纯合子时 $y=p^2$，子为杂合子时 $y=2pq$。p、q 各为随机人群提供生父、生母基因频率。

所以 PI 计算式为 $1/p^2$、$1/2pq$、$0.5/p^2$、$0.5/2pq$、$0.25/p^2$、$0.25/2pq$。PI 计算式见表 21－3。

3．单亲亲子鉴定

根据 PI 计算原则，将缺少的一方作随机人考虑进行计算，即 x 为假设父遗传生父基因和随机母遗传生母基因产生孩子的机会，y 为随机父和随机母生孩子的机会，即人群中孩子表现型频率。在极大多数基因组合的情况下，单亲亲子鉴定的 PI 值较双亲均检验的 PI 值小。

PI 计算可归纳简化为：

设 p、q 分别为随机人群生父基因频率，

（Ⅰ）父为纯合子时 x＝1，父为杂合子时 x＝0.5；

（Ⅱ）子为纯合子时 y＝p，子为杂合子时 y＝2p；

（Ⅲ）父、子为相同基因型杂合子时 x＝0.5（p＋q），y＝2pq。

STR 基因座 PI 简化计算式见表 21－4。

表 21－4　STR 基因座单亲 PI 简化计算

基因组合		x	y	PI
父	子			
AA	AA	1	p	1/p
AA	AB	1	2p	1/(2p)
AB	AA	0.5	p	1/(2p)
AB	AC	0.5	2p	1/(4p)
AB	AB	0.5(p+q)	2pq	1/(4p)+1/(4q)

4. 非经典的三联检验

有时会遇到一些案子，假设父因某种原因（如死亡）不能参与检验，而由一个与他有关系的人（R）（R 为假设父的亲属）代替假设父参与检验，亲子鉴定中的三联体的假设父不是真正的假设父，而是假设父的亲属 R。

LR 的两个统计检验假设：Hp 代表孩子的生父是一个与 R 有关的人，Hd 代表孩子的生父是一个与 R 无关的人。PI 或 LR 计算公式（见表 21－5）与经典三联体检验不同，经典三联体检验，Hp 代表假设父是生父，Hd 代表随机人是孩子的生父。

这时计算 PI 就需要考虑假设父与其亲属的同宗系数。同宗系数（coancestry coefficient，θ）表示由一个群体中的两人随机选出一基因为同源基因的概率。各种不同情况的 θ 值不同：父－子、同胞：$\theta=1/4$；祖父－孙子、同父异母同胞（半同胞）、叔－侄子：$\theta=1/8$；第一代表堂兄弟：$\theta=1/16$，第二代表堂兄弟 $\theta=1/64$。表 21－5 为假设父的亲属是生父时 PI 的计算。

表 21－5　PI 值简化计算表

表型			PI
孩子	母	假设父的亲属	
AA	AA	AA	$1/[p_i(1-2\theta)+2\theta]$
		AB	$1/\{2[p_i(1-2\theta)+\theta]\}$
		BC	0
	AB	AA	$1/[p_i(1-2\theta)+2\theta]$
		AB	$1/\{2[p_i(1-2\theta)+\theta]\}$
		BC	0
AB	AA	BB	$1/[p_j(1-2\theta)+2\theta]$
		BC	$1/\{2[p_j(1-2\theta)+\theta]\}$
		CD	0
	AB	AA	$1/[(p_i+p_j)(1-2\theta)+2\theta]$
		AB	$1/[(p_i+p_j)(1-2\theta)+2\theta]$
		BC	$1/\{2[(p_i+p_j)(1-2\theta)+2\theta]\}$

续表

		CD	0
	AC	BB_j	$1/[p_j(1-2\theta)+2\theta]$
		BD	$1/\{2[pj(1-2\theta)+\theta]\}$
		CD	0

5. 同父异母同胞检验

在一些涉及遗传继承案件中，某二个个体 X、Y 为不同母亲所生，个体 X 的父亲 Z 已亡，个体 Y 也声称是 Z 的儿子，因此要求亲子鉴定。由于 Z 已亡，无法参与检验，只能进行同胞间亲缘鉴定。

Hp：假设 X 和 Y 是同胞

Hd：假设 X 和 Y 不是同胞

似然比

$$LR=\frac{Pr(GX,GY\mid Hp)}{Pr(G_X,G_Y\mid Hd)}=\frac{Pr(G_X,G_Y\mid Hp)}{Pr(G_X\mid Hd)\,Pr(G_Y\mid Hd)}$$

表 21－6 为个体 X 和 Y 拥有各种基因型情况下的 LR 值，此时没有考虑近亲和人群结构因素。

表 21－6　遗传纠纷 LR 值计算

GX，GY	Pr（G_X，G_Y | Hp）	Pr（G_X | Hd）Pr（G_Y | Hd）	LR
A_iA_i，A_iA_i	$0.5p_i^3(1+p_i)$	p_i^4	$1/2+1/(2p_i)$
A_iA_i，A_iA_j	$0.5p_i^2p_j(1+p_i)$	$2p_i^3p_j$	$1/2+1/(4p_j)$
A_iA_i，A_jA_j	$0.5p_i^2p_j^2$	$p_i^2p_j^2$	1/2
A_iA_i，A_jA_k	$p_i^2p_jp_k$	$2p_i^2p_jp_k$	1/2
A_iA_j，A_iA_j	$0.5p_ip_j(4p_ip_j+p_i+p_j)$	$4p_i^2p_j^2$	$1/2+1/(8p_i)+1/(8p_j)$
A_iA_j，A_iA_k	$0.5p_ip_jp_k(4p_i+1)$	$4p_i^2p_jp_k$	$1/2+1/(8p_i)$
A_iA_j，A_kA_l	$2p_ip_jp_kp_l$	$4p_ip_jp_kp_l$	1/2

6. 突变存在时 PI 计算

在一些亲子鉴定中，多个基因座中只有一个基因座上出现不匹配，一般认为是由于遗传突变造成的，而不是由于亲子关系。当存在这种突变时，计算 PI 值时需要考虑突变率，计算公式如表 21－7。

表 21－7　突变存在情况下三联体与单亲的 PI 计算

父母参与的三联体		
孩子	母亲：假定父亲	LR
AA	AA:BB	$\frac{M\delta_{AB}(1+3\theta)}{2(2\theta+(1-\theta)p_A)}$
AA	AA:BC	$\frac{(M\delta_{AB}+M\delta_{AC})(1+3\theta)}{4(2\theta+(1-\theta)p_A)}$

续表

AA	AB:ZZ（Z≠A）	$\frac{M\delta_{Az}(1+3\theta)}{2(2\theta+(1-\theta)p_A)}$
AA	AB:WZ（Z，W≠A）	$\frac{(M\delta_{Aw}+M\delta_{Az})(1+3\theta)}{4(\theta+(1-\theta)p_A)}$
AB	AA 或 AC:ZZ（Z≠B）	$\frac{M\delta_{Bz}(1+3\theta)}{2(1-\theta)p_B}$
AB	AA 或 AC:WZ（W，Z≠A）	$\frac{(M\delta_{Bw}+M\delta_{Bz})(1+3\theta)}{4(1-\theta)p_B}$
AB	AB:CC	$\frac{(M\delta_{AC}+M\delta_{BC})(1+3\theta)}{2(2\theta+(1-\theta)(p_A-p_B))}$
AB	AB:CD	$\frac{(M\delta_{AC}+M\delta_{BC}+M\delta_{AD}+M\delta_{BD})(1+3\theta)}{4(2\theta+(1-\theta)(p_A+p_B))}$
	单亲鉴定	
孩子	假定父亲或母亲	LR
AA	BB	$\frac{M\delta_{AB}(1+2\theta)}{2(\theta+(1-\theta)p_A)}$
AA	BC	$\frac{(M\delta_{AB}+M\delta_{AC})(1+2\theta)}{4(\theta+(1-\theta)p_A)}$
AB	CC	$\frac{(M\delta_{AC}p_B+M\delta_{AB}p_A)(1+2\theta)}{4(1-\theta)p_Ap_B}$
AB	CD	$\frac{(M\delta_{AC}+M\delta_{BC}+M\delta_{AD}+M\delta_{BD})(1+3\theta)}{8(1-\theta)p_Ap_B}$

式中 $M\delta_{AB}$ 等为等位基因 B 突变为 A 的突变率，θ 为同宗系数

7．多基因座 DNA 指纹图

比较被检验人的多基因座 VNTR－DNA 指纹图，先确定母亲与孩子共有的片段，然后比较假定父与孩子的片段，确定不能在母亲中找到的孩子的片段是否能在假设父中找到，当多基因座 VNTR－DNA 指纹鉴定不排除亲生关系时，可用三种方式表示亲生关系相对机会（Jeffreys）：

（1）以 $1-X^n$ 计算值，X 表示随机两个体共有一个片段概率，n 表示假设父与子共有片段数：当 $1-X^n$ 计算值大于 0.9995，可以认定有亲生关系；这是实际案件检验中常用的方法；

（2）假设父/子共有片段概率：根据假设父与孩子之间共有片段多少来判断是否存在亲生关系。根据经验值，当共有片段率在 0～35% 为非父子关系，45%～100% 为亲子关系。适用于单亲检验。

（3）"陌生带"与非母带的比值：陌生带指的是不能在母亲和假设父图谱中找到的孩子片段，涉及突变等因素，有其特殊价值。当"陌生带"与非母带的比值在 0～0.2 范围内认定有亲生关系，0.2～0.4 不能下结论，0.4～1 排除亲生关系。

PI 理论值可以接近无穷大，RCP 理论值可非常接近 100%，但不能达到 100%。RCP 作为衡量亲子关系可能性大小的指标，按统计学标准，当 RCP 达到 95%，已具有肯定亲生关系的意义。但究竟达到多大，才能肯定亲生关系，国际上无统一标准。Ian w Evett 等

(1998) 提出的标准见表 21－8。目前实验室常用常染色体 STR 基因座，根据在真实家系中存在二个基因座的突变的可能，以及国际上多数实验室的标准（如荷兰免疫血液学红十字中心实验室等），肯定亲权关系的标准定为 RCP≥99.95%或 PI 值大于 2 000 比较妥当，当 PI＜1999，RCP＜99.95%时，应加测遗传标记以达到认定或否定的正确结论。

表 21－8 国际上常用标准

W 值	意义	W 值	意义
＞0.9973	可以肯定有亲子关系	0.80～0.5	不可能肯定是否有亲子关系
0.99～0.9973	极可能有亲子关系	0.1～0.5	倾向排除亲子关系
0.99～0.95	非常可能有亲子关系	0.05≤W≤0.1	不大可能有亲子关系
0.95～0.9	可能有亲子关系	0.0027≤W≤0.05	不可能有亲子关系
0.9～0.8	倾向有亲子关系	≤0.0027	可以排除亲子关系

21.3 亲子鉴定

通过 DNA 标记检测与分型，分析受检者之间的基因型是否符合遗传规律，从而判定有无亲生关系。亲代与子代的遗传基因型违反孟德尔定律（因突变影响需二个以上遗传标记），便可直接排除亲权关系。对于不能排除的亲子关系，需要进行对其亲子关系概率的计算。排除假三联，肯定真三联。

1. 标准的三联体亲子鉴定

由于母子关系已确定，需要鉴定假设父亲与孩子之间是否存在生物学亲子关系，可以用常染色体标记。比对母子的遗传标记，找出孩子的生父基因（obligatory gene），当假设父不具有生父基因，则可排除他与孩子的生物学父子关系；当假设父具有生父基因，则可能为生父，通过计算父权指数或父权相对机会对其为生父可能性的大小作出定量的估计，根据实验室评估标准，确定是否存在父子关系。

2. 女—子—男三联体亲子鉴定

由于父母与子女关系均不确定，需要鉴定父母双方与孩子的亲生关系，可以用常染色体标记。比对孩子与假设父、假设母的遗传标记，若孩子具有假设父、假设母所缺乏的基因，则排除他们之间有亲生关系；若孩子的基因可分别从假设父、假设母得到，则他们之间可能有亲生关系。根据双亲不定的模式计算出 PI 和 RCP 值，作出定量评估，判定亲生关系。

3. 单亲亲子鉴定

可采用：(1) 常染色体上标记，比较假设父（或假设母）与孩子的基因型，如果假设父（或假设母）的一对等位基因与孩子的一对等位基因完全不同，则可排除他们之间有亲生关系；假设父（或假设母）与孩子有相同基因时，他们之间可能有亲生关系，按单亲模式计算出 PI 和 RCP 值，作出定量评估，判定亲生关系。(2) 如果是鉴定母子（女）关系，可以采用母系遗传的线粒体 DNA 分析，如果两者分型不同，可以直接排除，如果相同，不能排除，它们之间可能有亲生关系，计算单倍型频率，是否达到认定水平。(3) 如果是鉴定父子关系，可以采用父系遗传的 Y－STR 分析，如果两者分型不同，可以直接排除，

如果相同，不能排除它们之间可能有亲生关系，计算单倍型频率，是否达到认定水平。

4. 隔代、同胞、旁系亲属间的亲子鉴定

由于某种原因（如死亡），被假设的父亲不能参加检验，而由假设父的亲属参加检验。或者父或/和母已亡，或因某种原因不能参与检验，如为了认亲、移民、继承财产、入户等，要求进行隔代、同胞或叔侄、姨甥等旁系人员间的亲权鉴定。情况将变得复杂些。因为假设父的基因型需要通过假设父的亲属来推断，有时由于亲属太少无法推断出假设父的基因型，此时不能凭常染色体的标记进行鉴定。比如，在一个家庭中，同胞之间基因组合有三种可能：① 一对等位基因完全相同；② 有一个等位基因相同；③ 一对等位基因完全不同。所以在无法推断假设父的基因型时，检验 STR 基因座不能提供任何排除信息。这类血缘关系鉴定要作出排除结论，利用 Y－STR 和 mtDNA 检测较具有优势，可以非常简单排除非亲缘关系。

根据亲子鉴定的种类与被检验人员的关系情况，确定检验的方法。

（1）同胞的血缘关系鉴定

同胞之间的血缘鉴定可分为以下两种情况：（1）只有 1 个同胞被检验，要求鉴定争议对象与他的血缘关系，简称为与单个人的同胞鉴定。（2）有 2 个或 2 个以上为确知有血缘关系的同胞被检验，要求鉴定争议对象与他们的血缘关系，简称为与多个人的同胞鉴定。

利用 Y－STR 和或 mtDNA 的遗传特点进行鉴定。如果受检者之间的 mtDNA 测序结果不同，或假设弟与兄之间的 Y－STR 分型结果不同，则可排除他们有同胞关系，如果相同，则可能有同胞关系。应用 Y－STR 和 mtDNA 标记，可以直接简单、明了地确定比对样品是否有亲缘关系。

前面已提到常染色体 STR 基因座各等位基因是共显性遗传，在一个家庭中，同胞之间基因组合可有三种可能，有时无法推断它们的同胞关系。

①与单个人的同胞鉴定　这类血缘鉴定要作出排除结论，只能利用 Y－STR 和 mtDNA 检测。由于常染色体上标记无法正确推断出生父基因，即使他们的基因型不同，也不能排除他们的同胞关系。

②与多个人的同胞鉴定　除了 Y－STR、mtDNA 检测可提供排除信息外，常染色体 STR 检测也能提供排除信息。如根据已知同胞的基因型（当两者基因型完全不相同时）可推出亲代的基因组合，计算遗传生父、生母基因的机会，然后在假设兄弟与亲代之间按照双亲皆疑的方法计算 PI 值。

同胞鉴定时应该注意：① 尽可能多检验已知血缘关系的同胞，以提供更多的信息来推出父母亲的基因型；② 检测更多的遗传标记系统，从已知同胞基因型推测父母基因型组合，计算 PI 值；③ 尽可能地利用 Y－STR 和 mtDNA 多态性在同胞鉴定中的作用，尤其是与单个人的同胞鉴定时，只检测常染色体 STR，可能得不出明确结论，一定要检测 Y－STR 和 mtDNA 多态性。

（2）旁系的血缘鉴定

旁系间叔侄、姨甥、堂表兄弟（姐妹）等等的血缘鉴定，与同胞鉴定一样，只有 Y－STR、mtDNA 检测才能提供直接的排除信息；若两个人的 Y－STR 分型不同，可排除他们之间有叔侄、堂兄弟关系，若两个人的 mtDNA 测序结果不同，可排除他们之间有姨甥、

表兄弟（姐妹）关系。在检测结果不违反遗传规律时，计算 PI。

（3）祖孙隔代（——三代）亲子鉴定 PI 计算

根据第一代基因型推测第二代可能的基因型，对于常染色体共显性遗传标记而言，最多可推出子代的 4 种基因型，每种概率为 0.25。如果孩子的生父基因不能在假设祖父、祖母的基因中找到，可以排除他们的子孙关系。如果能找到，则不能排除，计算 PI 值。首先依据第二代基因型计算传给第三代生父（或生母）基因的频率，纯合子传递生父（或生母）基因频率为 1，杂合子传递生父（或生母）基因频率为 0.5，计算生父（或生母）基因由第一代传递给第三代的频率。根据检验对象不同情况分别计算 PI：孩子母亲参与检验，则按仅怀疑父亲三联计算，PI 为生父基因由第一代传递给第三代的频率与随机人群生父基因频率之比（如上述例子）；孩子母亲未参与检验，如只有祖父、祖母、孩子参与检验，则计算生父基因频率后，按单亲计算 PI。

如果受检人为男性，也可以用 Y-STR 标记进行检验，如果单倍型不一致，可以直接排除他们的子孙关系，如果一致，他们可能来自同一父系，按单倍型计算亲权指数。

21.4 亲子鉴定应注意的问题

理论上，亲代与子代个体间的每一个遗传标记均应符合遗传规律，所以只要有一个遗传标记违反遗传规律就可否定亲生关系。但 DNA 基因座存在突变率高的缺陷，在亲子鉴定中孩子图谱出现陌生带或不符带，无法弄清是突变还是没有亲子关系出现的排除，故一般需有 2~3 个以上独立系统不符亲权关系作出的排除结论方为可靠。在鉴定中应注意：

（1）单基因座 DNA 纹印图　严格地说，单基因座 DNA 纹印图是表型图，不能认为是基因型，因为目前检测分离系统无法区分开只差一个重复单位的等位基因片段，因此当只检出一条片段时，有三种可能：真正纯合子；或因 DNA 纹印图中等位基因呈连续分布，片段较大，肉眼无法分辨两个片段等形成的假纯合子；或是因小片段电泳迁移出凝胶留下较长片段。故假设父与孩子均表现为纯合子时，一定要慎重，搞清一条带情况究竟属于哪一种情况，排除假纯合子现象，以免假排除。

（2）STR 系统　已发现在一些基因座如 D7S820，D21S1，D8S1179 引物结合位置发生突变，引物退火时无法结合，导致一个等位基因丢失，成为一个无效的等位基因，表现为一个假纯合子，使本来符合孟德尔遗传的两个体表现为违反遗传定律，造成假排除。这种假纯合子在 PCR 产量上有差异，带的强度与杂合子的带强度相当，比真正纯合子带的强度弱一倍。改变引物序列，可以克服这种情况。实验室最好拥有两套不同的试剂盒，可以互相印证。

（3）小卫星和微卫星突变　大量研究报道小卫星和微卫星基因座存在较高的突变率，小卫星的突变率比微卫星高，比如小卫星和微卫星突变率为 1.1×10^{-3}（$0.2-3.3 \times 10^{-3}$）与 2.0×10^{-3}（$0.9-4.1 \times 10^{-3}$），有些报道小卫星和微卫星突变率为 0.6×10^{-3}（$0.3-1.1 \times 10^{-3}$）与 1.4×10^{-3}（$1.0-1.9 \times 10^{-3}$）。研究发现，Y 染色体突变比常染色体高 2 倍，男性突变率比女性高 17 倍，父方突变比母方高，而且男性随着年龄增长单个重复单位突变增加，但女性中则没有这种年龄效应。

因此，亲子鉴定中，出现一个或二个基因座排除时，不能排除其亲子关系，需要进步

加测遗传标记，进步分型。出现突变时 PI 值计算与正常的 PI 值计算有所不同，要考虑到突变率。

(4) 对于小卫星 VNTR 基因座，由于较小等位基因优先扩增现象比较明显，要注意较小等位基因的丢失。

(5) 多基因座 DNA 指纹图系统，突变或甲基化现象，孩子的图谱中存在 1~2 条陌生带，已在 DNA 指纹图章节详细介绍，并不影响亲子关系的认定。

(6) 线粒体 DNA 分型要注意异质性的问题，亲代、子代或同一母系的相关个体之间存在一个碱基的差异，当出现单个碱基差异不能直接排除他们的亲缘关系。但异质性很少同时发生在二个碱基处。因此，本章中所述关于 mtDNA 测序结果不同，均指有二个或二个以上位置的碱基不同。

(7) 另外需注意输血及骨髓移植对亲子鉴定的影响，接受过输血的病人要待 2~3 个月后才能做亲子鉴定，以免供血者遗传标记的的影响。而接受骨髓或脐血移植的白血病病人，因供者血细胞已取代了受者血细胞，所以不再适宜以血液为标本做亲子鉴定。

由于突变或其他因素的存在，国际上对亲子鉴定的排除提出一些建议，主要是：

(1) 只有 1 个遗传标记不符合遗传规律，不能轻易作出排除结论，必须加测其它系统。随着检测项目的增加，对于不是生父的男子，必定还有其它遗传标记可排除亲子关系。若增加检测项目，排除遗传标记不再增加，则可考虑原来排除的那 1 个遗传标记是由突变或非典型遗传方式造成的，此时若亲子关系相对机会已超过 0.9995，则可作出认定结论。

③2 个 STR 基因座或同一染色体上的两个遗传标记不符合遗传规律，需慎重对待，宜加测遗传标记后再具体分析，加测标记数目要考虑 CCE 值达到 0.9997 以上，结论较为稳妥。

(3) 有 3 个及以上独立遗传关系（即非连锁）的遗传标记不符合遗传规律，则可作出排除亲子关系的结论。因为假设 DNA 单一基因座突变率为 0.002 ，三个位点同时突变造成的错误排除概率仅为 4×10^{-9}。

亲子鉴定是 DNA 分析检验中的一项重要内容，越来越多的案件涉及亲权鉴定，比如，某一杀妻案，罪犯将其妻子杀死并分尸，将妻子所用过的东西也清理干净，在被怀疑是分尸现场找到芝麻粒大小的血迹，要求鉴定此血迹是否是死者的血，由于没有死者比对样品，只能通过比对死者父母的 DNA，鉴定现场血迹与死者父母是否有亲子关系，来确定血迹是否是死者的血。这种通过亲子鉴定确定罪犯犯罪事实的案件日益增加，因此亲子鉴定应十分严谨，排除亲子关系要求有二个以上独立遗传标记的不符合遗传规律，对于认定亲子关系更要慎重，父权指数或亲权指数要达到一定的数值。亲子关系的认定是一个相对值，有的只是在已检验的遗传标记系统中符合遗传定律，随着遗传标记的增加，有可能会排除其亲子关系，因此要求一定高的父权指数，使从概率上说随机人群不可能拥有这种亲子关系。单亲检验时，各遗传标记系统的非父排除率比三联体检验的低，如 D18S51 基因座，三联非父排除率为 0.785，二联非父排除率为 0.563。如不违反遗传规律，单亲检验由于缺少母亲一方的遗传信息，计算的父权指数一般也较双亲均检验的低。为了得出明确的结论，故检测项目应比三联检验多，才能使 CCE 大于 0.9995，不排除亲子关系时，RCP

值能大于 0.9995。如只检测 STR 基因座，一般需检测 16 个以上基因座。

线粒体 DNA 因为是在核染色体外，遗传时不经过有丝分裂和减数分裂，通过卵细胞传递，所以母亲与孩子的 mtDNA 序列是相同的，在母 - 子单亲鉴定中有较大应用价值。其 PI = 1/ 某个 mtDNA 遗传标记的人群频率。但必须注意，与母亲有血缘关系的个体，无论男女均有与母亲相同的 mtDNA 型。当检验结果为母子之间有不同的 mtDNA 型时，可作出否定结论；有相同 mtDNA 型时，不能作肯定结论，尤其是当怀疑姐妹两人或姨甥女之间有一人为生母时，检测 mtDNA 不能提供任何信息，应该以常染色体 STR 为主。

由于父子具有相同的 Y 染色体遗传标记，所以检测 Y - STR 基因座鉴定血缘关系不需要母亲的任何信息，在父子单亲鉴定中有特殊意义。但 Y - STR 与 mtDNA 类似，与父亲有血缘关系的男性个体，均有相同的遗传标记，故在兄弟两人或叔侄之间确定谁是生父时，检验 Y - STR 无价值。

第二十二章　其他法医 DNA 分析应用

法医 DNA 技术的发展，在为人体的各种组织和体液材料进行个人识别和亲子鉴定提供了极为有效方法的同时，也被用于研究解决其他法医和生物物证鉴定问题。包括应用 DNA 技术去研究解决法医病理学、法医人类学的课题，以及应用 DNA 技术鉴定人以外的其他生物物证。现在在这些方面都已经取得了一定的进展。如根据 DNA 分析推断年龄段、推断受伤和死后间隔时间等方面有了些值得继续探讨的结果。在对家畜、宠物、野生动物、昆虫等的研究上取得了肯定的可应用的结果，有些已经试用于刑事和民事案件鉴定。在植物和微生物的 DNA 分析上也开始作了一些有价值的研究和应用尝试。所有这些，虽然都还有大量的研究工作要继续做，但都为扩大 DNA 技术的应用范围，扩大生物物证的检验范围，全面地利用各种生物物证，增加物证提供的信息种类和数量，提高其在侦察和破案中的线索价值和证据价值提供了可能。

22.1 DNA 检验应用于年龄推断

年龄鉴定是法医学检验的重要内容。通过生物检材准确估计个体年龄可以为案件的侦破提供重要线索。用 DNA 技术从分子水平上推断年龄，丰富了个体识别在年龄推断方面的内容，具有重要的法医学实用价值。目前用于年龄推断研究的 DNA 标记主要是线粒体 DNA 和端粒。

22.1.1 mtDNA 4977 bp 的缺失

Holt（1988）和 Wei（1992）发现骨骼肌肉及其他软组织的线粒体 DNA 的缺失情况与年龄有着显著的相关性。1997 年 Meiβner 及其同事提取不同年龄段（3 个月 ~ 97 岁）个体的线粒体 DNA，用定位于线粒体 DNA 8285/13928 核苷酸的一对引物 L35/H35 进行扩增，当线粒体 4977bp DNA 发生缺失时，这对引物可以扩增出片段长度为 667bp 的产物，如果没有缺失存在，就检测不到这 667bp 片段。实验发现，获得 667bp 扩增产物的模板 DNA 的量随被试验者年龄的增大而减少。在 0 ~ 20 岁人中，检测不到 4977bp 的缺失，在 20 ~ 30 岁人中，检测量为 1 000ng，而在 70 岁以上人群中，只要 1ng 总 DNA 就可以检测到 4977bp 的缺失（图 22 - 1，表 22 - 1）。4977bp 的缺失与个体的年龄有着密切的关系，缺失水平与年龄呈正相关，4977bp 的缺失可以作为推断年龄的有效标记。因此，如果只用 1ng DNA 检出 4977bp 的缺失，暗示着该个体的年龄在 50 岁以上，如果检测不到缺失，说明该个体年龄可能在 20 岁以下。

当然需要注意的是，线粒体 DNA 的缺失受内源和外源因素的影响，如 Parkinson 疾病、心肌病等、日光照射引起皮肤线粒体 DNA 的缺失等。但不管怎样，生理年龄与线粒体 DNA 的 4977bp 缺失有紧密的相关性，这在某些情况下，可能对推断个体年龄有用。

分析方法：

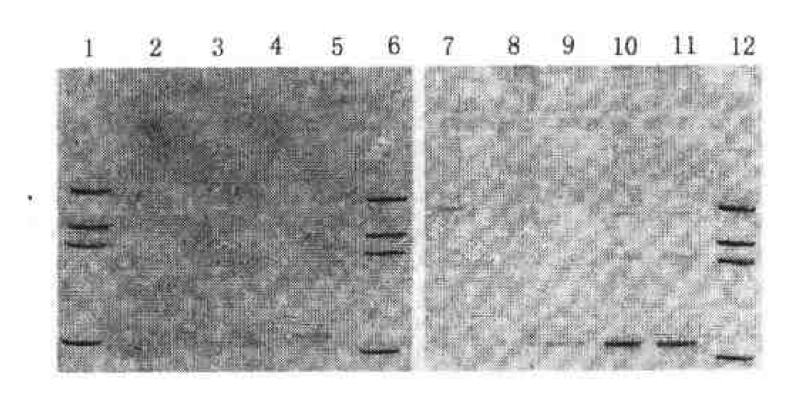


图 22－1 随年龄增加的 mtDNA 的 4977bp 的缺失的电泳银染结果

1，6 和 12 号为 phi×174/HaeⅢ的标准分子量；2：45 岁个体的 10ng 总 DNA 的结果；3：54 岁个体；4：67 岁个体；5：80 岁个体；7：37 岁个体；8：45 岁个体，；9：54 岁个体；10：67 岁个体

表 22－1 能检测出线粒体 DNA 的 4977bp 缺失的总 DNA 量与个体的年龄

小组	年龄（岁）	样品数	总 DNA 量（ng）						
			1 000	500	100	50	10	5	1
1	0－10	0	0	0	0	0	0	0	0
2	11－20	0	0	0	0	0	0	0	0
3	21－30	13	13	12	12	10	0	0	0
4	31－40	9	9	9	7	6	1	0	0
5	41－50	9	9	9	7	7	6	2	0
6	51－60	13	13	13	13	12	11	7	4
7	61－70	13	13	13	13	13	12	9	6
8	71－80	8	8	8	8	8	8	8	8
9	>80	11	11	11	11	11	11	11	11

1.PCR 引物序列

L15（3304nt）：5′－AAC ATA CCC ATG GCC AAC CTC CTA C －3′

H15（3836nt）：5′－ GGC AGG AGT AAT CAG AGG TGT TCT T－3′

L35：5′－CTC TAG AGC CCA CTG TAA AGC TAA C －3′

H35：5′－CTA GGG TAG AAT CCG AGT ATG TTG GAG AA －3′

引物 L15 和 H15 扩增 533bp 片段，L35 和 H35 扩增含 4977bp 缺失的 667bp 片段。

2．反应条件 在 25μl 反应体系中，含有 400 pmol/L 每个引物，200μmol/L dNTPs，1u Taq DNA 聚合酶，50mmol/L KCl，10mmol/L Tris－HCl（pH8.4），1.2mmol/L $MgCl_2$，一定量 DNA 模板。

3．循环参数 94℃40s，65℃40s，72℃50s，40 个循环，最后 72℃7min。

4．检测 7μl 产物在 7.5%C/2.5%T，60mmol/L 甲酰胺 PAG 中在 20W 下电泳 45～55min，以 Phi X174/HaeIII 的片段为标准分子量。银染显色。

22.1.2 端粒

端粒（telomere）是指真核细胞染色体末端的 DNA 序列，其生物学功能是保持染色体

的稳定。人类的端粒 DNA 主要由 TTAGGG 片段串联重复构成，其重复次数为 800～3000 次，长度为 5 至 15 kb。Harley 等人通过培养不同年龄（0～91 岁）和人体不同部位的成纤维细胞，观察端粒 DNA 长度的变化。发现随着培养细胞分裂次数的增多，端粒 DNA 长度逐渐缩短，端粒 DNA 的量也逐渐减少，供体年龄与端粒 DNA 长度及总量成明显负相关。这种变化与细胞的静止和生长状态无关，只与细胞的分裂次数相关。细胞平均倍增一次，端粒 DNA 的长度大约减少 50bp。Vaziri 检测人类外周血细胞的端粒 DNA 长度时，也发现了这种规律，端粒 DNA 长度每年丢失 41 ± 7.7 bp。端粒 DNA 丢失的速率在人的整个生长过程中是不相同的。青年时期端粒 DNA 的丢失较缓慢，随着年龄的增大，特别是进入老年后，端粒 DNA 长度由于不完全复制而逐渐缩短。不同种族、不同性别、不同组织端粒 DNA 有差异，端粒 DNA 长度不完全相同。这些发现提示，通过选择合适人群和组织，检测端粒 DNA 长度和端粒 DNA 量的变化，可以发现端粒 DNA 变化与年龄的关系。这种关系有可能用于法医生物学检材推断年龄的研究。

22.2 DNA 分析在死亡时间推断上的应用

准确推断死后间隔时间（postmortem intreval，PMI）是法医病理学研究的重要内容。目前多用尸温的下降速度与死后时间的曲线关系来推断死亡时间。但该方法影响因素较多，实际应用受到许多条件限制。因此，寻找其他方法十分必要。研究表明细胞核内 DNA 在死后即开始自溶性降解，3 天内受环境温度及死亡方式的影响较小。因此研究组织 DNA 含量变化规律与死亡时间的关系有重要价值。应用流式细胞仪分析组织中含降解 DNA 细胞数与含完整 DNA 细胞数的比例，显示其比值随 PMI 延长而增加。提示这种关系有可能用于推断死亡时间。

22.3 昆虫 DNA 分析在法医鉴定中的应用

昆虫学对法医学尸体死亡时间的推断起着重要作用。早在宋慈的《洗冤录》中就介绍了根据物证上的昆虫确定罪犯，法国 M. Bergeret（1855）对巴黎郊区一所房子的壁炉内的骷髅进行检查，根据尸体上的残留昆虫表明死亡时间在几年前，论证凶手是该房子的原主人，而不是现在的房主。昆虫帮助法医学家推断死亡时间、推断案发现场等。

利用昆虫学进行尸体死亡时间推断的理论基础主要是根据与尸体有关的昆虫的生长发育的形态变化。各种昆虫在不同的生长时间其形态特征不同，利用形态学来判断昆虫的发育阶段的最大难题是大多数昆虫的幼虫的形态相似，由于昆虫种类繁多，很难判断是哪一种昆虫。而且还必须将现场采集的幼虫拿到实验室进行培养至成虫，然后才能进行准确的鉴定，这样，势必延迟了数天，甚至数个星期。更为重要的是，在培养过程中，有些幼虫死了，或者有时在现场提取到的只是昆虫的一些残存物，已无法根据形态学进行鉴定。但是，实际工作中要求在现场采集后能马上进行快速、准确可靠的鉴定。此外也应用生化技术进行昆虫种类的鉴定。同工酶（alloenzyme）是应用最广泛的标记，但这些标记常常只是在生活史的某一个特定时期得以表达，而且容易被降解。象 ELISA 等这样灵敏的免疫学技术对于区分近缘种类的昆虫还是不够灵敏。

DNA 分析在生物物证鉴定中的应用，为昆虫的种类鉴定提供了一个有效的方法。以

DNA 为基础的分析方法可以分析任何阶段的昆虫，大大缩短时间。线粒体 DNA 是此类鉴定最合适的标记，因为线粒体 DNA 分子小，相对耐降解，每个细胞内含有数千个拷贝，突变率高，即使是近缘种类的昆虫也存在较大的序列差异，可以区分开来。对于昆虫种类鉴定，目前主要有两大类方法与标记，一是对 mtDNA 细胞色素氧化酶进行测序分析，或用限制性内切酶分析 CO-I 的 PCR 产物，二是用各种引物进行 RAPD 分析。

DNA 技术除用于进行昆虫种属鉴定外，还可以根据从食人血和尸体的昆虫嗉囊分泌物中提取到人 DNA 的分析，揭示昆虫与有关人或尸体的关系，这有时可能有助于案件的侦破。

22.3.1 昆虫线粒体 DNA 细胞色素氧化酶序列分析

每种昆虫细胞都含有线粒体，线粒体 DNA 中细胞色素氧化酶亚基Ⅰ（cytooxidase subunit gene I，CO-I）在不同品种昆虫间存在较大的碱基序列差异。Sperling（1994）分析了三种常见丽蝇科昆虫，即伏蝇（Phormia regina）、丝光绿蝇（Phaenicia sericata）和亮绿蝇（Luxilia illustris）线粒体 DNA 的 2.3kb 序列，结果发现丝光绿蝇和亮绿蝇之间有 118 个核苷酸序列不同，伏蝇和亮绿蝇间有 186 个核苷酸序列不同，伏蝇和丝光绿蝇之间有 192 个核苷酸序列不同，根据这些序列差异，轻而易举地将三种幼虫区分开来。即使昆虫已经死亡，仍然可以在一天内做出鉴定。选用合适的限制性内切酶，用 PCR-RFLP 分析CO-I基因，也可以区分开这三种昆虫。Vincent（2000）也用测序方法分析昆虫线粒体 DNA 的细胞色素氧化酶亚基 I（CO-I），进行欧洲 BLOW 蝇 的鉴定。Vincent 用引物 CO-I 2f 和 CO-I 3r（序列同 Sperling，1994）扩增测序模板 DNA 片段，用 CO-I 2f 或 CO-I 3r 进行测序分析Calliphoridae家族的7个欧洲不同种类的昆虫（图22-2），不同种类昆虫之间存在较大的差异，差异度在 2.9%~13.9%范围（表 22-2）。亲缘关系最近的 Lucilia caesar（Linne）和 Lucilia illustris（Meigen），其差异度为 2.9%，亲缘关系最远的 Lucilia caesar（Linne）和 Protophormia terraenova（R.-D），差异度为 13.9%。这些种类之间的差异是由于碱基取代造成的，没有插入或缺失。

表 22-2　不同种类昆虫在 mtDNA CO-I 的差异和取代

	L. sericata (Meigen)	L. illustris (Meigen)	L. illustria (Meigen)	C. vicina (r.-D.)	C. vomitotria (Linne)	P. Terranovae (R.-D.)	C. macellaria (Fabricius)
L. sericata (Meigen)	/	10.2%	8.6%	9.5%	10.2%	10.9%	8.0%
L. illustris (Meigen)	14	/	2.9%	9.5%	10.9%	10.9%	9.5%
L. caesar (Linne)	12	4	/	9.5%	10.2%	9.5%	9.5%
C. vicina (R.-D.)	13	13	15	/	5.1%	13.1%	11.7%
C. vomitoria (Linne)	14	15	16	7	/	13.1%	10.9%
P. Terranovae (R.-D.)	15	15	16	18	19	/	10.9%
C. macellaria (Fabricius)	11	13	13	16	15	15	/

不同的材料，如地上发现的死成虫，酒精保存的蛆、蛹和空壳，室温下干燥的空壳

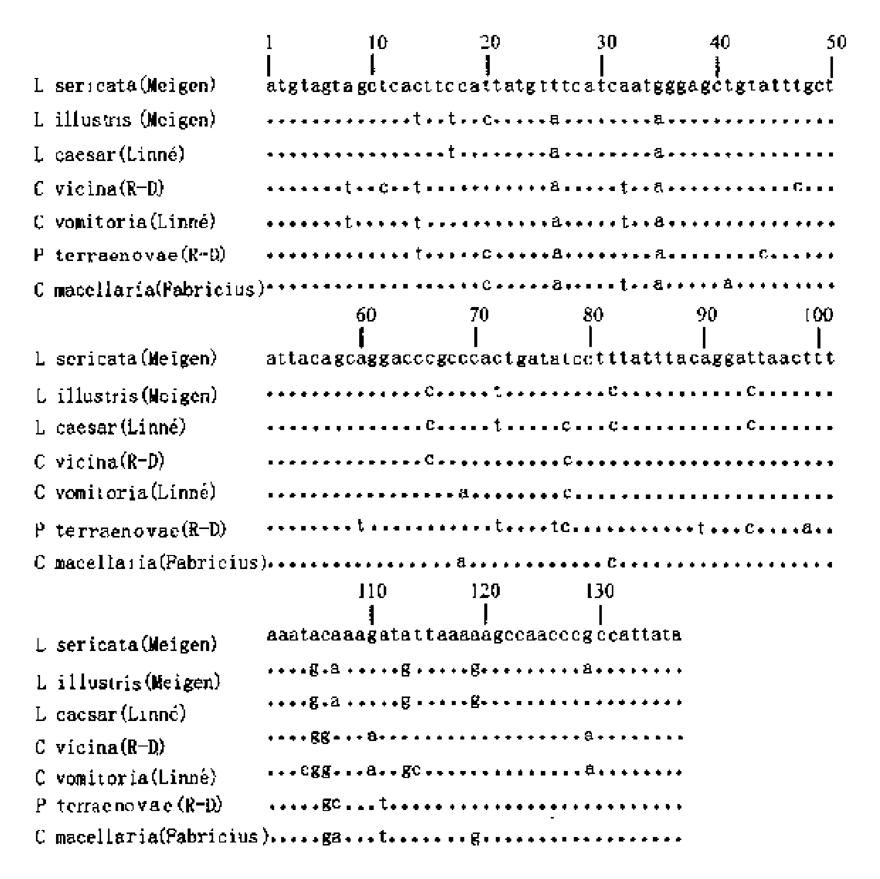


图 22-2　mtDNA 的 CO-I 的序列比较

等，均得到了分析结果，因此序列分析 mtDNA 的 CO-I 可以进行幼虫、保存较差的昆虫和空蛹的种类鉴定。

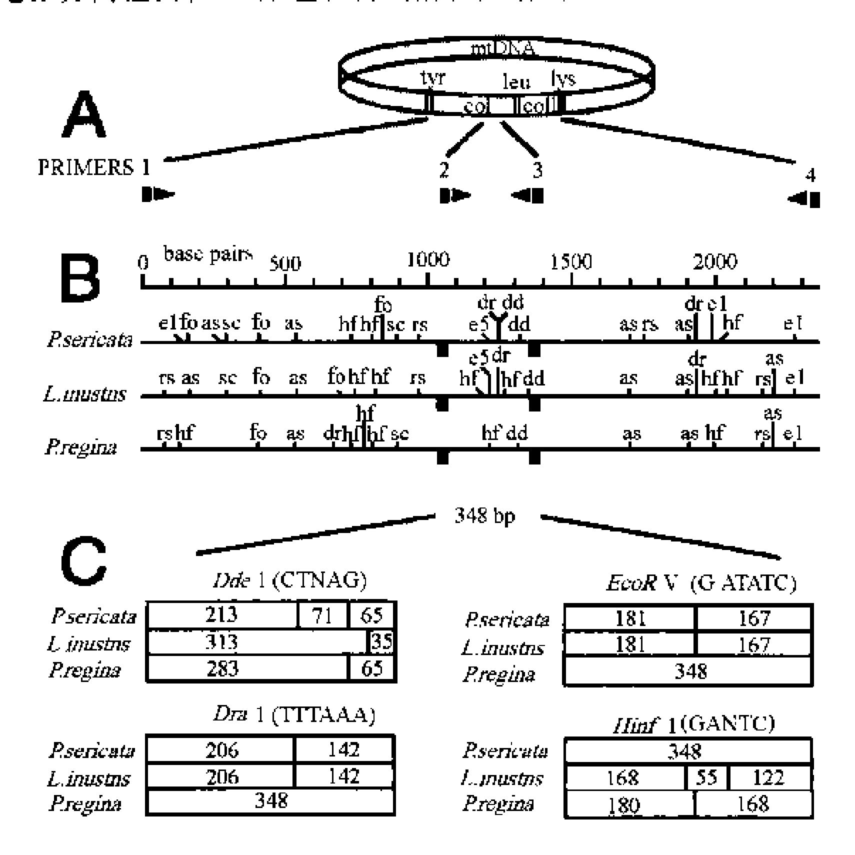
分析方法

采用改进的 Harrison（1987）DNA 提取法提取成蝇基因组总 DNA。其步骤为：（1）将成蝇胸部放入 1.5ml 离心管内，置于液氮中磨碎成粉状；（2）加入 800μl 的 Lifton 缓冲液（0.1mol/L Tris 缓冲液，0.2mol/L 蔗糖，0.05mol/L EDTA，0.5%SDS，pH9.0），裂解细胞膜和抑制 DNA 酶；（3）轻轻旋涡振荡匀浆液，室温放置 15min 至 2h；（4）加入 120μl 的 8mol/L 乙酸钾，混匀后置冰浴 15min 至 3h，离心 15min；（4）上清液用等体积酚抽提之后，再用等体积氯仿/异戊醇（24:1）抽提；（5）加等体积异丙醇，-20℃下放置 25min 沉淀 DNA。（6）用 500μl 70%乙醇洗 DNA 沉淀，干燥后，DNA 沉淀悬浮于 200μl TE 缓冲液中。

乙醇保存标本和干燥标本的 DNA 提取除了以下几点不同外，其他同成蝇胸部 DNA 提取：（1）乙醇保存样品在液氮中研磨前先真空干燥除去乙醇；（2）在加入 Lifton 缓冲液后立即加 12μg 蛋白酶 K，60℃消化 2h。（3）两次酚抽提以保证彻底除去蛋白酶 K；（4）DNA 沉淀悬浮于 100μl TE 缓冲液中。

按图 22-3 中引物序列进行扩增测序，反应条件是：在 50μl 反应体系中，含有 1μl DNA 溶液，95℃3min，45℃2min，72℃1.5min，一个循环后，94℃1min，45℃1min，72℃1.5min，循环 30~35 次，最后 94℃1min，45℃ 2min，72℃5min，一个循环。双链 PCR 产物经 Centricon 100 纯化后，用末端标记法测序。

扩增产物中存在 9 个酶切点，选用 9 种限制性内切酶进行 PCR 产物酶切分析（图 22－3）。大约 1μg PCR 扩增产物在 15μl 反应体系中进行酶切反应。消化产物用 1.5%～2% 琼脂糖凝胶电泳分离检测，EB 染色，得到酶切图谱（图 22－4）。



Dde I (CTNAG)

P.sericata	213	71	65
L.inustus	313		35
P.regina	283		65

EcoR V (G ATATC)

P.sericata	181	167
L.inustus	181	167
P.regina	348	

Dra I (TTTAAA)

P.sericata	206	142
L.inustus	206	142
P.regina	348	

Hinf I (GANTC)

P.sericata	348		
L.inustus	168	55	122
P.regina	180	168	

图 22－3 3 种丽蝇线粒体 DNA 扩增区和引物位置、9 个限制性内切酶切点位置

A：线粒体 DNA 扩增区和引物位置

引物 1：5′－TAC AAT TTA TCG CCT CAC CC－3′

引物 2：5′－CAG CTA CTT TAT GAG CTT TAG G－3′（CO－I 2f）

引物 3：5′－CAT TTC AAG C/TT GTG TAA GCA TC－3′（CO－I 3r）

引物 4：5′－GAG ACC ATT ACT TGC TTT CAG TCA TCT－3′

B：整个序列中的限制性内切酶切点：as = Ase，dd = GdE，dr = Dra，el = EcoRI，e5 = EcoR V，fo = Fok I，hf = Hinf I，rs = Rsa I，sc = Sac I

C：酶切引物 2 和 3 扩增细胞色素 C 氧化酶亚基 I 基因（COI）的 348bp 扩增产物

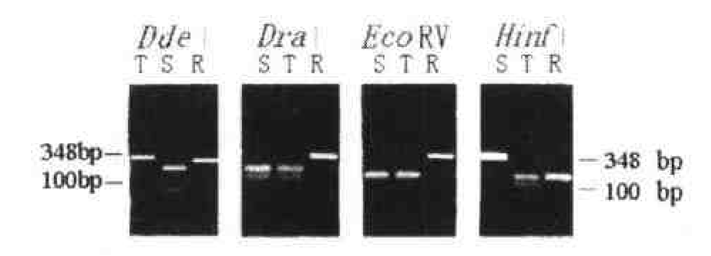


图 22－4 4 种限制性内切酶酶切产物电泳结果

用 4 种内切酶（Dde I，Dra I，EcoR V，Hinf I）将伏蝇与其他两种蝇区分开来，

Dde I 和 Hinf I 将丝光绿蝇与亮绿蝇分开。事实上，Hinf I 可以把这 3 种蝇区分开。

22.3.2 昆虫的RAPD分析

除了分析线粒体DNA CO-I外，还可以用RAPD方法进行昆虫种类鉴定。

Benecke（1998）采集腐尸上的昆虫，用9个引物（表22-3）对双翅目和鞘翅目（diptera，coleoptera）昆虫进行9个体系RAPD分析，并对该方法进行了有效性的研究。用这种方法鉴定了一起案件中涉及的装尸体袋里面和外面的蛆。

表22-3 RAPD引物

引物名称	序列（5′-3′）	退火温度（℃）
ERIC-2	XAAF TAA GTG ACT GGG GTG AGC	66.4
DAF4	XCGG CAG CGC C	
REP1R	[a]XACG TCG ICA TCA GGC	
REP2	XACG TCT TAT CAG GCC TAC	
PRIMER1	XGGT GCG GGA A	
PRIMER2	XGTT TCG TCC	
PRIMER3	XGTA GAC CCG T	
PRIMER4	XAAG AGC CCG T	
PRIMER5	XAAC GCG CAA C	
PRIMER6	XCCC GTC AGC A	
BERG1254	XCCG CAG CCA A	

X：荧光染料

分析方法

在25μl反应体系中，含有1.5u Taq DNA聚合酶，10mmol/L Tris-HCl（pH9.0），50mmol/L KCl，1.5mmol/L $MgCl_2$，200μmol/L dNTPs，25pmol/L引物，BSA，50ng DNA。扩增循环：95℃预变性5min，95℃1min，36℃1min，72℃2min，循环45次。PCR产物在5.5%PAG上电泳，荧光检测。用每一个引物进行RAPD分析，每个昆虫均能得到许多条带，而且用一个引物可以区分不同种类的昆虫，根据已知标本，可以确定待测样品的种类。

22.3.3 昆虫中的人DNA分析

一些食血或食腐尸的昆虫在饱食后，需要一定的时间消化所采的食物，在一段时间内它们不再采食，从它们的嗉囊中可以提取到被食人的DNA。例如，食血为生的蚊子，吸完血后，蚊子落在墙上、顶棚或窗帘等地方，在吸血后2~3天才消化掉所吸的血，产卵后才开始吸新的血。Coulson（1990）用同位素标记探针杂交，在蚊子吸食20h后的蚊子血中检测到人的DNA，Sato（1993）用非同位素探针进行斑点杂交检测，在蚊子吸食100h后，在蚊子中检测到人的DNA。

Kreike和Kanpfer（1999）用PCR技术分析，蚊子吸食26h后得到与人一致的D1S80和HLADQα分型。他们早上将有4个人同住过的房间内的蚊子打死，装入50ml的Facon管中，室温放置26小时后，冰冻起来。用标准方法提取DNA，再用Chroma spin-100柱纯化。然后进行D1S80和HLADQα分型。结果在其中4只蚊子中提取到人的DNA，得到了D1S80和HLADQα分型（表22-4）。其PCR扩增效率相当于含有0.5μg的人基因组DNA。

表 22-4 蚊子在 HLADQα 和 D1S80 基因座的分析结果

昆虫/人	HLA-DQα	D1S80	血来自于
蚊子#1	1.1/1.2	n.d.	A
蚊子#2	1.1/1.2*/3/4	T24/T37	A+D
蚊子#3	1.1/1.2	T24/T24	A
蚊子#4	4/4	T18/T18	B
个体 A	1.1/1.2	T24/T24	-
个体	4/4	T18/T18	-
个体	1.1/4	T19/T19	
个体	3/4	T24/T37	-

n.d.：表示没有检测到

寄生于人的虱子排泄物中也含有人 DNA。Reploy（1994）将数种虱子放在一些自愿受试者的皮肤上进行试验，结果在这些虱子排泄物中提取到高达 10ng 的人 DNA。Lord（1998）报道了从寄生人身体的虱子中提取到人 DNA，分析了 8 个虱子，8 个虱子均得到了人 DNA，并且与寄主的 DNA 序列完全一致。

Wells（2001）实验，他采集德克萨斯州 San Antonio 附近的野苍蝇的卵，进行孵化。将孵化的幼虫移到一内有人肝脏的玻璃缸内。待幼虫长到大约 12mm 长时将其中一部份幼虫置于在 -20℃的 70%乙醇中保存，而将另一些幼虫继续长成蛹，变成成虫。从形态学确定为蓝瓶子苍蝇（bluebottle fly cynomyopsis cadaverina）。支解每个幼虫，将嗉囊切成 3mm 长，用 QIAamp 柱纯化，进行 mtDNA HV2 测序分析和苍蝇 mtDNA CO-I 序列分析。结果发现，从嗉囊中提取到的 DNA 拥有与寄主（肝脏）一致的人线粒体 DNA 的 HV2 序列，只是来自嗉囊 DNA 测序图谱较对照的血 DNA 稍微难读，但在成年的苍蝇中未发现。

实验结果表明，昆虫的嗉囊中存在有人 DNA，昆虫嗉囊和肠内容物是检测人 DNA 的最理想材料。分析昆虫的嗉囊有可能确定昆虫来自于何处，与案件是否关联，有助案件的分析。例如，性虐待和身体残害案件中，受害人和罪犯同在一室住过一定时间，或一些抢劫团伙，在某一旅馆同一房间内住过数日。由于门、窗紧闭，蚊子不能飞出或飞入房间。在他们走后短时间内，采集分析室内蚊子中的人 DNA 有可能成为一项有意义的证据。

22.4 动物 DNA 分析在法医鉴定中的应用

法医 DNA 分析技术除了可对人类进行 DNA 分析外，还可对宠物、家畜、野生动物等进行种属鉴定、个体识别及亲权关系鉴定。事实上，动物的毛、骨骼等很早就是法医鉴定的检验材料之一。随着法制的加强，侦审对证据的重视和科技的发展，已经有越来越多的刑事案件和民事案件涉及此类鉴定。

用于动物鉴定的 DNA 技术主要有线粒体 DNA 的测序分析、RAPD 分析和 STR 等分析。

22.4.1 狗和猫的 DNA 分析

狗和猫是人类生活中接触最多的动物。狗可能作为进攻人的肇事者而受到检验，或者在犯罪分子身上或物品上发现有狗的毛，意欲通过检验这些狗毛和可疑狗的毛来判定他们

的相关性而被送到实验室来。匈牙利刑事技术研究所用 DNA 技术进行狗的个体识别和亲子鉴定，并受理此类案件。Aaspollu（2001）报道在爱沙尼亚的一个小村子里发现一具年轻妇女的尸体，法医鉴定此妇女受到狗的攻击。在妇女的外衣上收集到狗的毛和唾液斑，用线粒体和 STR 对住在附近的 6 只狗与现场提到的毛与唾液进行检验比对，确定了嫌疑狗与案件的关系。

1. 狗线粒体 DNA 测序分析

Savolainen（1997）利用 nested 扩增相当于人 mtDNA H16403 – L15996 的 319bp 片段作为测序模板 DNA，对狗毛进行测序分析。测序分析 52 个品种 102 只家狗的 mt DNA HVI 序列（257bp），发现有 23 个突变点，形成 19 个单倍型，21 个位置是转换，一个位置同时有转换和颠换的发生，另一个位置（单倍型 10）发生 1bp 的插入，而在单倍型 13 含有一 67bp 的串联重复，单倍型频率为 20.6% ~ 1%，最少见的 3 个型存在于 1% 的个体中，最常见的 6 个型存在于 75% 个体中，排除率为 0.88（图 22 – 5、图 22 – 6）。未观察到这些单倍型有明显的种系特异性，但是某些单倍型在一些品系中出现较高。如在各种品系的 15 只猎犬中，有 9 只都是单倍型 4，占 60%，在 6 只拉布拉多猎狗（Labrador Retriever）中，5 只为单倍型 4，占 83%，所有阿富汗猎犬（Afghan houndss）（3 只）都为单倍型 6，在 7 只德国牧羊犬中，其中 4 只是单倍型 5，占 57%；单倍型 8 只在 4 只狗观察到，它们属于与北斯堪的纳维亚（northern Scandinavian）种系的 jamthund 和挪威猎犬（Norwegian elkhound，Grey）。唯一一只切萨皮克湾猎犬（Chesapeake Bay Retrievder）含有唯一的单倍型 13 序列，它含有一个从 101 ~ 167 位置核苷酸的 67bp 串联重复的插入。该犬的母狗及 3 个同胞含有同样的序列，单倍型 13 除了这个插入外，其他序列与单倍型 4 相同。因此，单倍型 13 的序列可能来自于单倍型 4。

10	20	30	40	50	
CTGAGATTCT	TCTTAAACTA	TTCCCTGACA	CCCCTAATT	CATATATTGA	50
ATTACCCCTA	CTGTGCTATG	TCAGTATCTC	CAGGTAACC	CTTCTCCCCT	100
CCCCTATGTA	CGTCGTGCAT	TAATGGTTTG	CCCCATGCAT	ATAAGCATGT	150
ACATAATATT	ATATCCTTAC	ATAGGACATA	TTAACTCAAT	CTCATAGTTC	200
ACTGATCTAT	CAACAGTAAT	CGAATGCATA	TCACTTAGTC	CAATAAGGGC	250
TTAATCA					257

图 22 – 5　狗的 mtDNA 高度区 I 序列

该序列作为与其他单倍型比对的标准序列，下划线碱基表示下图 22 – 6 中的多态核苷酸可能位置。此序列的第一个核苷酸和最后一个核苷酸位置相当于人 mtDNA 的 15966 和 16402 位置。

单倍型	多态性核苷酸位置																							
5	22	30.1	53	78	96	103	123	155	(162)	165	181	182	190	195	197	202	204	206	209	213	220	222	223	
1	G	T	–	T	C	C	C	A	A	T	C	T	T	T	T	G	C	C	T	A	A	T	G	A
2	.	.	.	C	.	.		.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
3	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.
4	.	.	.	C	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.
5	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	T	.	.	.	.
6	A	.	.	C	.	T	.	.	.	.	T	.	C	.	.	A	.	T	.	G	G	.	A	.
7	A	.	.	C	T	T	.	.	.	.	.	C	.	.	.	A	.	.	.	G	.	C	.	.
8	A	.	.	C	.	.	.	.	.	.	.	.	.	.	C	A	T	A	C	G	G	.	.	.
9	.	.	.	C	.	.			.	.	.	.	.	.	.		T	.	.	.	.	.	.	.
10	A	.	C	C	.	T	.	.	.	.	.	.	C	.	.	A	T	.	.	G	G	.	A	.
11	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.
12	.	.	.	C	.	.	.	.	G	.	.	.	.	.	.	A	.	.	.	T	.	.	.	.
13*	.	.	.	C	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.
14	.	.	.	C	.	.	.	G	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.
15	.	C	.	C	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	T	.	.	.	.
16	.	.	.	C	.	.	.	.	.	.	.	.	.	.	C	A	.	.	.	T	.	.	.	.
17	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	T	.	.	.	G
18	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.
19	A	.	.	C	T	T	T	.	.	.	.	C	.	.	.	A	.	.	.	G	.	C	.	.

图 22－6 在 102 只犬中发现的 19 个序列的 23 个多态性核苷酸位置。

核苷酸位置根据上图 22－5 参照序列 1 定位的。只有一种序列变化较大。从引物 D4 的 3′端开始记数。在单倍型 10 中发现一个参照序列 1 中没有的核苷酸的插入，用小数点表示。用“－”表示参照序列中相应的缺失，单倍型 13 用“＊”标记，它含有一个 67bp 的串联重复的插入。

单倍型	品系	数量	频率%
1	Border collie	2	9.8
	Irish Setter	2	
	Chow – chow	1	
	Collie	1	
	Leonberger	1	
	Norweigian Buhund	1	
	Samoyede	1	
	Crossbreed	1	
2	Chow – Chow	1	2.9
	Tibetan Terrier	1	
	Chinese crested	1	
3	Rottweiler	2	15.7
	English Setter	2	
	Norw. Elkhound, Grey	2	
	Border Terrier	2	
	Papillon	2	
	Whippet	1	
	Springer spaniel	1	
	Poodle	1	
	Icelandic Sheepdog	1	
	Tox Terrier	1	
	Japanese Spitz	1	
	Siberian Huskey	1	
4	Labrador Returever	5	20.6
	Golden retriever	2	
	Flat – coated Retriever	2	
	Norfolk Terrier	2	
	Papillon	1	
	Newfoundland	1	
	Alaskan Huskey	1	
	Giant Schnauzer	1	
	Boxer	1	
	Leonberger	1	
	Kuvasz	1	
	Schipperke	1	
	German Shepherd dog	1	
	Crossbreed	1	
5	German Shepherd dog	4	14.7
	English Setter	2	
	Dachshund, Wirehaird	2	
	Samoyede	1	
	Border Collie	1	
	Old English sheepdog	1	
	Keeshond	1	
	Hamiltonstovare	1	
6	Crossbreed	1	8.8
	Afghan	3	
	Golden Retrievee	2	
	Groenendael	1	
	Irish Water Spaniel	1	
	Otter hound	1	
	Tibetan Spaniel	1	
7	Giant Schnauzer	2	6.9
	Alaskan huskey	1	
	German Shepherd dog	1	
	Airedal Terrier	1	
	Jamthund	1	
	West highl. White	1	
	Terrier	1	
8	Jamthund	2	3.9
	Norw. Elkhound, Grey	1	
	Crossbreed (Elkhound grandmother)		
9	Irish Setter	1	3.9
	S; t Bernhard	1	
	Greyhound	1	
	Tibetan Terrier	1	
10	Flat – coated Retriever	1	2.0
	achshund, Wirehaired	1	
11	Irish Wolfhound	2	2.9
	Pyrenean Mastiff	1	
12	Labrador Retriever	1	1.0
13	Chesapeak Bay Retriever	1	1.0
14	Fox Terrier, Wire	1	1.0
15	Golden Retriever	1	1.0
16	Norwegian Lundehund	1	1.0
17	Cav. King charles Spaniel	1	1.0
18	Siberian Huskey	1	1.0
19	German Shepherd dog	1	1.0
单倍型	品系	数量	频率%

图 22－7　分析的 102 只狗情况及其它们的序列变异体分布

方法：

DNA 提取：方法同人类毛发

PCR 扩增与测序

引物 D1 和 D2 为外圈 PCR 引物，D3 和 D4 为内圈 PCR 引物及测序引物，D3、D4 的 3′末端位置相当于人 mtDNA H16403 与 L15996。D3 的 5′端含 M13 - 21 序列，使测序反应可以进行引物标记法测序，D4 标有生物素，主要用来纯化 DNA 片段。

D1：5′ - AGA GGG ACA TTA CGA GCA AGG - 3′

D2：5′ - CCT AAG ACT TCA AGG AAG AAC C - 3′

D3：5 - ′TGT AAA ACG ACG GCC AGT TTG ATG GTT TCT CGA GGC ATG G - 3′

D4：5′生物素 - CTC CAC CAT CAG CAC CCA AAG - 3′

第一次（外圈）PCR 反应　在 50μl 反应体系中，含有 5 ~ 10μl 模板 DNA，10mmol/L Tris - HCl（pH8.3），50mmol/L KCl，1.5mmol/L $MgCl_2$，0.1% Tween20，2μg/ml BSA，0.2mmol/L dNTPs，0.1μmol/L 引物 D1 和 D2，1u Taq DNA 聚合酶。扩增循环参数为：94℃预变性 2min，然后进行 94℃15s，65℃30s，72℃1min，循环 25 次。

第二次（内圈）PCR 反应　在 50μl 反应体系中，含有 2μl 第一次 PCR 扩增产物作为模板 DNA，10mmoL/L Tris - HCl（pH8.3），50mmol/L KCl，1.5mmol/L $MgCl_2$，0.1% Tween20，2μg/ml BSA，0.2mmol/L dNTPs，0.1μmol/L 引物 D3 和 D4，1u Taq DNA 聚合酶。扩增循环参数为：94℃预变性 2min，然后进行 94℃15s，69℃30s，72℃1min，循环 25 次，最后 72℃10min。第二次扩增产物纯化后作为测序模板，用 D3 或 D4 引物进行测序。

Savolainen（1998）又对该方法进行了改进，设计新的引物使测序模板变短，以便能分析各种检材。他采用了 semi - nested 扩增 125bp 片段作为测序模板 DNA。

引物序列：

D5：5′ - TGA TGG TTT CTC GAG GCA TGG - 3′

D8：5′ - CCC CAT GCA TAT AAG CAT GTA C - 3′

D7：5′ - ATT AAG CCC TTA TTG GAC TAA GTG - 3′

引物 D5 和 D8 为第一次 PCR 扩增引物，其扩增产物片段为 148bp，D7 和 D8 为 semi - nested 的第二次 PCR 扩增的引物，扩增产物片段为 125bp。

第一次 PCR 反应　在 50μl 反应体系中，含有 5 ~ 10μl 模板 DNA，10mmoL/L Tris - HCl（pH8.3），50mmol/L KCl，1.5mmol/L $MgCl_2$，0.1% Tween20，2μg/ml BSA，0.2mmol/L dNTPs，0.1μmol/L 引物 D5 和 D8，2u Taq DNA 聚合酶。扩增循环参数为：94℃预变性 2min，然后进行 94℃15s，66℃30s，72℃1min，循环 25 次。

第二次 PCR 反应　在 50μl 反应体系中，含有 2μl 第一次 PCR 扩增产物作为模板 DNA，10mmoL/L Tris - HCl（pH8.3），50mmol/L KCl，1.5mmol/L $MgCl_2$，0.1% Tween20，2μg/ml BSA，0.2mmol/L dNTPs，0.1μmol/L 引物 D7 和 D8，1u Taq DNA 聚合酶。扩增循环参数为：94℃预变性 2min，然后进行 94℃15s，65℃30s，72℃1min，循环 25 次，最后 72℃10min。第二次扩增产物纯化后作为测序模板，用 D7 或 D8 引物进行双向测序。

Savolainen（1998）用这种方法分析了几起涉及狗毛分析的案件。案例 1，一名妇女被谋杀，在尸体旁发现几根狗毛。过了没多久，又发现一名妇女被谋杀，同样的在尸体旁发

现数根狗毛。作案方式相似，都在现场留下形态上相似的狗毛，认为此两案为同一人所为。线粒体分析表明二个现场的狗毛的序列相同，均为单倍型 8（图 22－8），是非常少见的单倍型，采集 7 个可疑嫌疑人家的狗毛，进行线粒体 DNA 分析，它们的单倍型分别为 1，3，4，5，5，5，5 型，表明现场的狗毛与它们的不同，更为重要的是，现场发现的狗毛单倍型只在 3 个狗种中发现，缩小了现场遗留狗毛的狗范围。

案例 2，一个 4 人团伙抢劫了数家银行，在侦破中将其与另一起从一家养狗人家偷车并案。在逮捕时，从其中一个嫌疑人身上发现几根狗毛，为了调查这狗毛是否来自于丢车那家的狗身上，对狗毛进行了 mtDNA 分析，发现从狗身上拔下的狗毛与嫌疑人身上找到的狗毛均为单倍型 5（图 22－8），因此不能排除它们来自于同一个个体。

案例	样品	多态位置														单倍型	频率数据（%）
		10	13	29	30	38	43	45	50	52	54	57	61	68	70		
	标准序列	T	C	T	T	T	T		G	C	G	T	A	A	T	G	1
1	物证 1	–	–	–	–	–	C		T	A	C	G	G	–	–	8	3.9
	物证 2	–	–	–	–	–	C		T	A	C	G	G	–	–	8	3.9
	对照 1	–	–	–	–	–	–	–	–	–	–	–	–	–	–	1	12.7
	对照 2	–	–	–	–	–	–	A	–	–	–	–	–	–	–	3	16.7
	对照 3	–	–	–	–	C	–	–	–	–	–	–	–	–	–	4	21.6
	对照 4	–	–	–	–	–	–		–	–	–	T	–	–	–	5	15.7
	对照 5	–	–	–	–	–	–		–	–	–	T	–	–	–	5	15.7
	对照 6	–	–	–	–	–	–		–	–	–	T	–	–	–	5	15.7
	对照 7	–	–	–	–	–	–		–	–	–	T	–	–	–	5	15.7
2	物证 1	–	–	–	–	–	–		–	–	–	T	–	–	–	5	15.7
	对照 1	–	–	–	–	–	–		–	–	–	T	–	–	–	5	15.7

图 22－8　二起案件狗毛样品的 mtDNA 序列分析结果

以 Savolainenn（1997）报道的序列作为标准序列进行比较，序列从引物 D8 的 3′端开始

2. 狗和猫的线粒体 DNA 的串联重复序列分析

物种间线粒体 DNA 的结构非常相似，高度多态的控制区都有一个中央的保守区和旁侧的二个高变区（HV1，HV2）。猫和狗的线粒体 DNA 在控制区含有串联重复序列，猫的控制区大约有 1560bp，二个明显的串联重复序列位于中央保守区二个边缘区域，第一个序列的重复主型长度为 80～82bp，第二个则较短，6～8bp 重复单位。狗的控制区为 1313bp，只有一个串联重复序列，重复主型为 10bp（图 22－9）。

用引物 L029 和 H606 可以同时扩增猫和狗序列，进行串联重复序列的检测，每个个体得到多条带（图 22－10），每条带代表一个串联重复等位基因，个体之间有较大的差异，需要注意的是，个体内也存在较小的差异，同一根毛的不同长度部位，图谱有所不同。这一点，如同在人的线粒体 D－环测序分析时所观察到的异质性。

分型方法

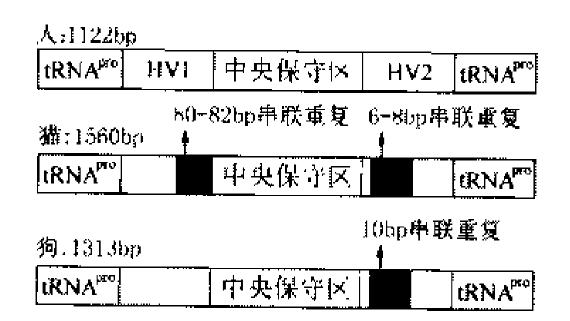


图 22-9 人、猫和狗的线粒体 DNA 位于 tRNA^pro和 tRNA^phe间的控制区结构

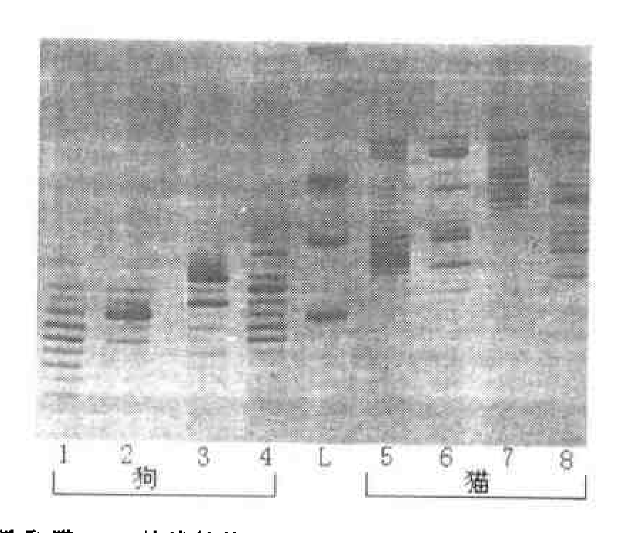


图 22-10 狗和猫 DNA 的线粒体 DNA 的串联重复序列的扩增产物电泳图谱

样品 1-4 为 4 只不同的狗，样品 5-8 为 4 只不同猫，L 为分子量标准，从上到下分别是 517bp，396bp，350bp

引物 L029 5′-TTC AGT CAA TGG TCA CAG GAC-3′

H606 5′-GGC GTA AGG GGG GTT TC-3′

反应体系 在 50μl 反应体系中，含有 0.5μmol/L 引物，0.2mol/L dNTPs，1.5mmol/L $MgCl_2$，1×PCR 反应缓冲液，0.2mg/ml BSA，1u Taq DNA 聚合酶，1μl DNA。

热循环参数 94℃45s，57℃45s，72℃3min，35 个循环。

产物检测 PCR 产物在 9.4% T，3.5%C 的聚丙烯酰胺凝胶中电泳分离，银染检测。

3. 应用 STR 分析进行狗的个体识别与亲子鉴定

关于狗的 STR 基因座，已有大量文献报道，有数百个 STR 基因座被研究，它们大多为二核苷酸重复单位，通常用来进行狗的亲子鉴定，具有较高的非父排除率。Koshinen (1999) 报道了用 3 个复合体系分析 10 个 STR 基因座（表 22-5，表 22-6)，并调查了 4 个品种的无关个体狗样品，在这 4 个品种中 10 个 STR 基因座的联合非父排除率分别为：拉布拉多猎狗（Labrador Retriever）99.78%，德国牧羊犬（German Shepherd）99.34%，金色猎犬（Golden Retiever）99.46%，芬兰猎犬（Finnish Hound）99.93%。选用合适的 STR 基因座，可以进行与人一样的个体识别与亲子鉴定。

表 22－5　STR 基因座的排除率

基因座	拉布拉多猎狗	德国牧羊犬	金色猎犬	芬兰猎犬
C04107B	0.14	0.45	0.12	0.30
CXX2054	0.5	0.38	0.39	0.53
CXX2004	0.22	0.28	0.29	0.24
CXX2001	0.24	0.50	0.34	0.51
CXX2132	0.74	0.45	0.50	0.65
CXX2137	0.58	0.54	0.37	0.69
CXX2146	0.54	0.40	0.71	0.66
CXX2088	0.28	0.30	0.27	0.33
CXX2168	0.55	0.48	0.40	0.36
CXX2175	0.38	0.04	0.54	0.62
联合	0.9978	0.9934	0.9946	0.9993

表 22－6　复合体系、引物用量与荧光标记

基因座	复合体系	量（pmol）	荧光标记
C04107B	1	3.5	FAM
CXX2054	1	1.2	FAM
CXX2004	1	1.8	HEX
CXX2001	2	5.0	TET
CXX2132	2	2.5	TET
CXX2137	2	2.5	HEX
CXX2146	2	20.0	HEX
CXX2088	3	1.5	HEX
CXX2168	3	6.0	TET
CXX2175	3	14.0	FAM

PCR 反应在 25μl 体系中进行，DNA 聚合酶量为 1.5u。热循环参数为 94℃3min 热变性，然后 94℃30s，60℃1min，72℃1min，35 个循环，最后 75℃15min。

Francisco（1996）用分子克隆方法，从基因库中筛选出 50 个四核苷酸重复 STR 基因座，表 22－7 列出了 47 个 STR 标记的克隆名称、引物序列、退火温度、多态性指数及片段大小。

表 22－7　四核苷酸标记的特点

克隆名称	重复序列	PIC 值	PCR 产物大小(bp)	退火温度	正链引物序列(5′－3′)	反链引物序列(5′－3′)
2001	(GATA)$_{8}$	0.62	131	58	TCCTCCTCTTCTTTCCATTGG	TGAACAGAGTTAAGGATAGACACC
2004	(GAAA)$_{13}$	0.79	237	58	CTAAGTGGGCAGCCTCCTCT	ACTCTGACCTACTCAGGTTGCA
2006	(GAAT)$_{7}$	0.48	191	58	TGGGGGCGTTAAGAGTAATG	CTAGGCCTAAACCCCTGAGC
2010	(ATGA)$_{10}$	0.69	228	59	AAATGGAACACTTCAGCATGC	CCCCTTACAGCTTCATTTTCC
2016	(GAAA)$_{25}$	0.76	318	58	CATTTTTTAAGGATGGAGACAGC	AACAGTGTCCCATGGCCTAC
2017	(GCTA)$_{6}$(GATA)$_{12}$	0.44	268	58	GTATCGAGTCCCCGCGTTC	CCCAGTACCACCTTCAGGA
2018	(AGA)$_{7}$(AATA)$_{12}$	0.58	149	58	TTGCTTTACCGCTTCATTTATT	GAGCCTGCTTCTCCCTCTG
2050	(TAAA)$_{5}$N(AAAC)5	0.42	258	59	GTATCGAGTCCCGCGTTG	CTGGGAGGTCCAAGGTTTAA
2054	(GATA)$_{16}$	0.66	151	58	GCCTTATTCATTGCAGTTAGGG	ATGCTGAGTTTTGAACTTTCCC
2060	(GAAT)$_{7}$	0.37	222	59	GTTTTGAGGAAGCCTTGCTG	CAAGCAACGGCCCAGTATTC
2062	(GAAT)$_{9}$	0.48	136	58	GGCTTCTGGAGACAGGCAT	CAGAACGCTGTCTAGCCCTT
2079	(GGAT)$_{47}$	0.55	275	58	CAGCCGAGCACATGCTTT	ATTGATTCTGATATGCCCAGC
2087a	(GAAA)$_{50}$	0.84	～300	58	CTGCCACATTCACTGATGC	CAACTCCCTCCCTCATTTCA
2087b	(GAAA)$_{15}$	0.76	～240	58	CTGCCACATTCACTGATCC	CAACTCCCTCCCTCATTTCA
2088	(TTTA)$_{10}$(TTCA)$_{4}$	0.59	117	58	CCCTCTGCCTACATCTCTGC	TACGGCATGCATATAACCACC
2096	(GAAT)$_{9}$	0.37	104	58	CCGTCTAAGAGCCTCCCAG	GACAAGGTTTCCTGGTTCCA
2097	(GAAA)$_{16}$	0.70	288	58	CAATGTCGAATTCCATGGTG	ATCCAGCAAGATGTGTTTGTG
2100	(GATT)$_{8}$	0.26	166	58	AGTCCCACATCAGGCTTCC	TCATCTAGCTGAGCTCTTCTCA
2107	(GAAA)$_{66}$	0.66	373	58	CATCATTATTGGGCTCCATG	CTGTGTCATCATTAAGCTAAGTCCT
2109	(GATT)$_{8}$	0.60	176	58	CAATCCAGCAACCCTCATCT	CAGGCATTGAGTCCCACATC
2119	(GAAAA)$_{19b}$	0.64	187	58	AGTTGTCAGCATTGTTTTCTCA	TTCAGTCTGGTSTGCAATGTCC
2130	(GAAA)$_{11}$	0.79	300	58	GCTGTCCTCCACTTTTCCTC	GTTAAGGAATAGTTCGGGGTCC
2131	(GAAA)$_{44}$	0.79	372	58	ATGAAGCCTCACGCCAAG	TGATCACACTCATCTCCCCA
2132	(GAAA)$_{35}$	0.85	272	60	CACTGGGAGTGGAGACTGCT	TGCACAGCCAACTAGAGGTG
2137	(GAAA)$_{21}$	0.86	185	58	GCAGTCCCTTATTCCAACATG	CCCCAAGTTTTGCATCTGTT
2138	(GAAA)$_{28}$	0.86	271	58	AATGTGCCCAACATTCCACT	AAGTCCCATGTAGGCTCC
2140	(GAAA)$_{18}$	0.77	146	58	GGGGAAGCCATTTTTAAAGC	TGACCCTCTGGCATCRAGGA
2141	(GAAA)$_{37}$	0.89	540	58	CAATGCTCCTTAGATTGCAGG	ACCCAACCAAACTACAAGCAAG
2142	(GAAA)$_{23}$	0.64	294	58	GCTCACATGTGTGCATGC	GCACTCCATGAGGCAGGT
2144	(GAAA)$_{14}$	0.63	289	58	GTGGCTCTTTTTGATGAGGG	CCTGGGTGGTTCAGTCAGTT
2145	(GAAA)$_{8}$	0.86	280	58	CAGTCGGGTTTAGGGTCTCA	CAGTCGGGTTTAGGGTCTCA
2146	(GAAA)$_{86}$	0.85	378	58	GGGCTCCACTAAGCACTATCT	TACCATAACAAAGCCAAGCC
2148	(GAAA)$_{21}$	0.72	480	58	GGCAGAAATGGAGACTATCAGG	GTCCCACGATGGACTCCC
2149	(GAAA)$_{53}$	0.76	308	58	CATCGAGCCTCCTTATCCAT	AACCCCTTAGCCACCCAG
2152	(GAAA)$_{55}$	0.89	433	58	ACCCCTCTTTCATGCTCTCTG	ACCATATCACTAGCGAAGGCAG
2155	(GAAA)$_{47}$	0.70	457	58	TGTAGATGATGGAGACATTGGC	AGGCAAATATGCCAAGGATG
2158	(GAAA)$_{44}$	0.86	274	58	ATGGCCACATCACCCTAGTC	CTCTCTCTGCATCRCRCATGAA
2159	(GAAA)$_{17}$	0.80	182	58	GAATCCCACATGGGGCTC	ATTAAGTTTTGAAAGCCAGGTAAG
2161	(GAAA)$_{21}$	0.70	250	58	TCAGCAAGAAACCCTCCACT	CATTCCCAACGGAGGACTCT
2162	(GAAA)$_{10}$	0.34	208	58	AAATTGAGCCCCACATCAAG	GCCATCTGGTACTGGTTTTTG
2164	(GAAA)$_{25}$	0.76	303	58	GATTATGACTCGAACCAAAGGC	TGGAGGAAGTTCATTAAGCAGC
2165	(GAAA)$_{43}$	0.91	326	58	GGATGGAGTCCCACATCG	CCTCACCCCAAATTCTTTAGC
2168	(GAAA)$_{20}$	0.75	213	58	GCAAATTACTTACTTCACTATGCC	TTGCAAGACTTCAACATGGC
2171	(GAAA)$_{73}$	0.82	～644	58	CTATGCCACTCTAGATTCTGGG	ACACATGCAGATAGCTCCTGG
2174	(GAAA)$_{16}$	0.33	201	58	CACCTGTTCTCATAGAATGCAG	AAGTCTCGCCTCGGGGTC
2175	(GAAA)$_{18}$	0.78	251	58	TTCATTGATTTCTCCATTCGC	AGCACTCTAAAAACTTCCCTCC
2201	(GAAA)$_{24}$	0.75	154	58	ATCAACAATGCATGCCACAT	GAGAAAAATAAATGCAAGCCC

22.4.2 马的 DNA 分析

15 个 STR 基因座用于马的个体识别和亲子鉴定（表 22－8，9），它们分为 3 个复合体系进行分析。

表 22－8　用于马的亲子鉴定 STR 标记

基因座	染料	等位基因数	M 等位基因长度（bp）	等位基因长度范围	参考文献
VHL20	FAM	10	97	89－107	
HTG4	FAM	8	133	127－141	
AHT4	FAM	14	158	138－170	
HMS7	FAM	10	179	167－189	
Lex3	FAM	13	210	194－220	
CA425	FAM	11	244	230－250	
HTG6	Joe	9	96	84－106	
AHT5	Joe	11	138	128－152	
HMS6	Joe	81	163	153－171	
ASB2	Joe	13	244	22－256	
HTG10	Tamra	15	101	89－117	
HTG7	Tamra	5	126	120－130	
HMS3	Tamra	11	160	150－174	
CA412	Tamra	9	196	188－208	
HMS2	Tamra	8	228	218－232	

表 22－9　马的 STR 基因座的复合 PCR 反应体系

复合体系	基因座	PCR 缓冲液(μl)	1.25mmol/L dNTPs(μl)	引物混合液*(μl)	Taq DNA 聚合酶(μl)	水(μl)	总量(μl)
8－plex	VHL20, HTG4, AHT4, HMS7, HTG6, HMS6, HTG7, HMS3	1.5	4.0	8	0.5	－	14
4－plex	HMS2, AHT5, ASB2, HTG7	1.5	4.0	4	0.5	4	14
3－plex	CA425, CA412, Lex3	1.5	4.0	3	0.5	5	14

* 0.5μl 10pmol/L 引物在反应前预先混合在一起

* 使用的酶为 AmpliTaq GOLD（5u/μl，PE）

分析方法

按表 22－9 准备 PCR 反应混合液，每个反应管中分装 14μl，再将 1μl 模板 DNA（20－50μg）加入到反应管中。95℃5min，然后 95℃1min，60℃1.5min，72℃2min，循环 34 次，72℃7min，4℃保存。1μl 扩增产物在 6% 变性聚丙烯酰胺凝胶中电泳检测、进行基因分

型。

22.4.3 猪的 DNA 分析

已有研究报道应用 STR 进行猪个体识别与亲子鉴定。Yue（1999）报道了用 8 个 PCR 复合体系扩增分布在 4 条染色体上的 29 个二聚、三聚核苷酸 STR 基因座。引物序列可在网站 http：//www.ri.bbsrc.ac.uk/pigmap/pigbase/pigbase 找到。

复合体系 1（S0094，Sw38，S0143，S0144，S0090 和 S0146）：

在 12.5μl 反应体系中，含有 100μmol/L dNTPs，50mmol/L KCl，10mmol/L Tris－HCl（pH8.3），1.5mmol/L $MgCl_2$，1.25u Taq DNA 聚合酶，50ng 基因座 DNA，分别为 0.2μmol/L，0.2μmol/L，0.2μmol/L，0.2μmol/L，0.6μmol/L，0.8μmol/L 的 S0064，Sw38，S0143，S0144，S0090 和 S0146 基因座引物。

复合体系 2（S0025，S0066，S0035）：

反应体系同复合体系 1，只是 Taq DNA 聚合酶改为 0.75u，引物浓度分别为 S0025 0.16μmol/L，S0066 0.4μmol/L，S0035 0.12μmol/L。

复合体系 3（S0282，Sw864，S0115，S0068）：

反应体系同复合体系 1，只是 Taq DNA 聚合酶改为 1.0u，引物浓度分别为：S0025 基因座 0.3μmol/L，Sw864 基因座 0.2μmol/L，S0115 基因座 0.06μmol/L，S0068 基因座 0.12μmol/L。

复合体系 4（S0605，S0076，S0212）：

反应体系同复合体系 1，只是 $MgCl_2$ 浓度改为 3.0mmol/L，引物浓度分别为：S0605 基因座 0.12μmol/L，S0076 基因座 0.18μmol/L，S0212 基因座 0.2μmol/L。

复合体系 5（Sw1021，Sw1057，Sw874）：

反应成分同复合体系 3，引物浓度分别为：Sw1021 基因座 0.2μmol/L，Sw1057 基因座 0.2μmol/L，Sw874 基因座 0.2μmol/L。

复合体系 6（Sw2001，w824，Sw581）：

反应成分同复合体系 1，不过反应体积改为 6.25μl，0.5u Taq DNA 聚合酶。引物浓度分别为：Sw2001 基因座 0.2μmol/L，w824 基因座 0.4μmol/L，Sw581 基因座 0.2μmol/L。

上述 6 个复合体系循环条件为：94℃45s，55℃30s，72℃45s，33 个循环。

复合体系 7（Sw520，S0102，S0147，TNFB）：

反应体系同复合体系 6，引物浓度分别为：Sw520 基因座 0.2μmol/L，S0102 基因座 0.4μmol/L，S0147 基因座 0.3μmol/L，TNFB 基因座 0.1μmol/L。循环条件：94℃45s，50℃ 30s，72℃45s，30 个循环

复合体系 8（Sw1329，S0215，S0083）：

反应体系同复合体系 6，引物浓度分别为：Sw1329 基因座 0.2μmol/L，S0215 基因座 0.2μmol/L，S0083 基因座 0.24μmol/L。循环参数：94℃45s，58℃30s，72℃45s，33 个循环。

PCR 产物分离检测　3μlPCR 产物在 5% 聚丙烯酰胺凝胶中电泳；1500V，45W，50℃，0.6×TBE 为电极缓冲液。根据分子量内标测定片段大小，进行基因分型。

Ruyter(1994)报道了 5 个 STR 基因座,这 5 个基因座的名称和引物序列等如表 22－10 所示。

表 22-10　5 个 STR 基因座引物序列

名称	重复子	引物（5′-3′）	等位基因数	杂合度
S0111	$(TG)_{16}$	A：TCA GTT ATT TCT GGC TAT CAT CTC B：TTG ATG TAG ACC ACC CAG CTA GTG	5	69%
S0112	$(TG)_{10}$ $(AG)_7$GG $(AG)_5$	A：AAT CCT GAG TAT CCG TAA TCA GGC B：TTG ACA TGA TGC AGA GAA GGA GTC	7	78%
S0113	$(TG)_6$ $(TA)_5N_{10}$	A：AGC CTC CGT GTA ATA TAA TCC TTC B：AGG ACA TCT CTC ATT CTC TGG CAG	5	72%
S0114	$(TG)_4$TA $(TG)_8$	A：CTC ACA AAA CGG GAT CTG TCC AGT B：CTG TGT GCC ATT CTT ATC AAA GCA	5	70%
S0115	$(GT)_{12}$ $(GA)_{21}$	A：TGA TGC ACT GTG TGG GCC ACA CCA B：ACC ATG GCT TGA GCT TGA GCC AGC	9	80%

分析方法：

在 20μl 反应体系中进行 PCR 反应，反应中含有 25ng 基因组 DNA，10pmol 引物，200μmol/L dNTPs，0.1u Taq DNA 聚合酶，1.5mmol/L $MgCl_2$，50mmol/L KCl，10mmol/L Tris-HCl（pH9.0），0.01%白明胶，0.1% Triton X-100。PCR 循环为：94℃1min，55℃30s，72℃1min，30 个循环，72℃10min。PCR 产物用 6%变性聚丙烯酰胺凝胶电泳分离，检测。

22.5 植物和微生物 DNA 分析在法医鉴定中的应用

犯罪分子在作案过程中无意间使自己身上、衣物、鞋袜、作案工具或交通工具携带有犯罪现场或和周围环境中的物体，如泥土、植物的叶片、花瓣、种子、花粉及果实或果汁等，或者尸体身上粘有这些物质。对这些物质进行分析，判定犯罪嫌疑人是否去过案件现场，或者确定第一犯罪现场位置，常常对一些案件的侦破具有重要作用。

对于案件现场及周围环境中这些物证，或根据植物的形态学特征或植物的分类学特征，或者根据土壤的土质、色泽、矿物质等特征加以区分、比较来进行判定。这要求检验人员具有丰富的知识与经验、熟练的技术。而且，在实际案件中，犯罪分子身上或尸体上附着的植物或泥土样品条件很差，凭借普通形态分类学方法可能无法确认和区分它们，因此对这些植物和泥土样品的检验与利用较少。随着分子生物学技术的发展，可以应用 DNA 技术对这些样品进行分析和检验，确定它们的种类与周围环境中同类物质的相似性。

22.5.1 应用 RAPD 技术进行植物鉴定

植物的品系鉴定与个体识别，常用的方法是 RAPD，曾被美国应用于一起刑事案件侦破中。在 1992 年 5 月 3 日早晨，两个晨起慢跑的人经过卡特彼勒公司在亚利桑那沙漠的一个废弃试验场时，发现了一具女尸。经警察调查，他们找到了一名叫马克·博根的卡车司机。但博根对警察说，他已有 15 年没有去过那个试验场附近了。他承认让死者搭过车，并和她发生过性关系，甚至那晚还和她打过架。警察搜查博根卡车时，在博根的卡车底座发现了两个假紫荆树的种子壳，谋杀现场正好有一颗假紫荆树被折断。为了证明该种子壳来自于现场被折断的那株植物，调查人员将 2 个假紫荆树的种子壳和那里的所有假紫荆树豆荚送往亚利桑那大学植物基因学家蒂莫西·海伦贯特里斯那里，进行 DNA 检验。他用

RAPD 技术检验所有的豆荚，结果发现卡车上发现的种子壳与现场被折断的那棵假紫荆树的 DNA 图谱一致，而与其他 3 组 41 棵假紫荆树的不同，为断案提供了有说服力的间接证据。

22.5.2 应用 DNA 技术分析土壤中微生物

新西兰环境科学研究所 Horswell（2001）分析土壤中微生物的核糖体基因（rDNA）确定土壤是否一致。其原理是土壤中含有大量微生物，包括真菌和细菌等，不同地方的土壤中所含的微生物种类有所不同，即使同一地方不同地表的土壤，其所含微生物也不同，用 DNA 技术分析土壤中微生物的 rDNA 图谱，比较二者的图谱，观察 DNA 峰或谱带是否一致，计算二者的相似性，确认二者土壤是否有关系。相似性用相似指数（similiarity index，SI）来表达。

$$SI = \frac{x}{2N}$$

其中 x 为二个电泳图谱中匹配的峰（带）数目，N 为二个图谱中所有出现的峰（带）数目。

图 22 - 11 是一起案件中涉及的靴子上的泥土与犯罪现场留下的靴子印的泥土中微生物的 DNA 分析结果，二者图谱非常相似，计算相似指数为 0.91，具有比其他地方的泥土高的相似指数，可以确定该靴子上的泥土与现场靴子印泥土一致，帮助办案人员确定穿用此靴子的人曾到过犯罪现场。

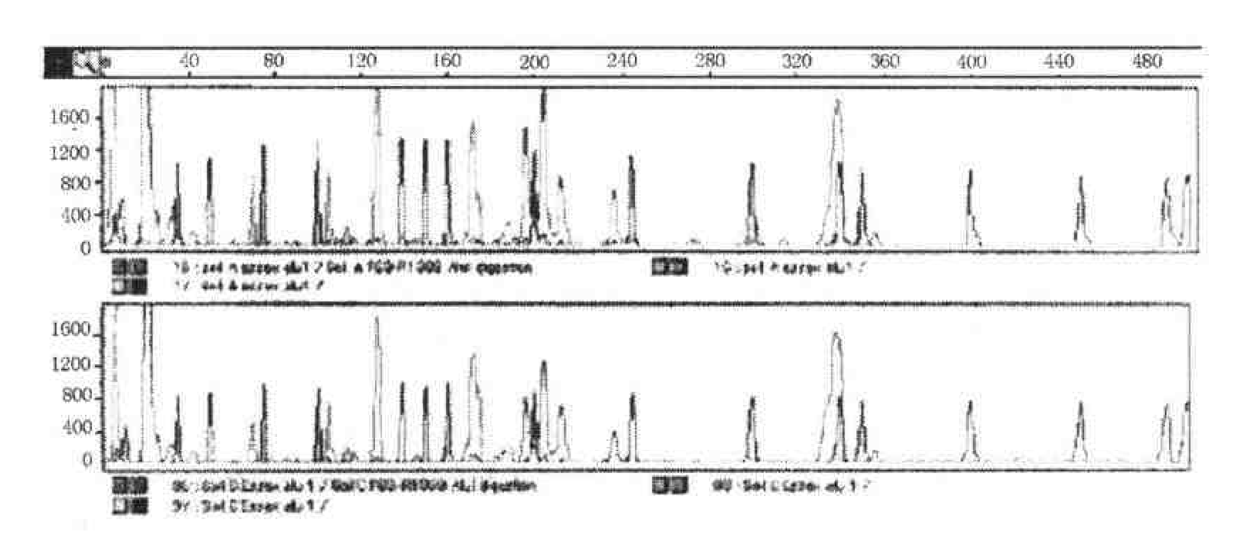


图 22 - 11 一起案件中涉及的靴子上的泥土与犯罪现场留下的靴子印泥土中微生物 DNA 分析结果

第二十三章　质量控制和质量保证与数据库

23.1 质量保证与质量控制

23.1.1 概述

DNA分型技术的应用使法医DNA分析人员能够对各种生物物证进行个人同一认定，对案件的侦破和审理起着十分重要的作用。因此，实验结果的高度准确性和可重复性十分重要，为了保证检验结果的可信性，必须制定科学的质量保障体系，对检验程序进行严格的控制。建立详细、严格的质量保证体系有助于建立科学、合理的分析程序。

质量保证指的是保证质量控制的正确实施，即保证DNA分型符合专门标准所采取的措施。DNA实验室应用的质量控制措施是保证DNA分型的质量符合和满足一定的标准。质量保证是一个动态的、需要定期检查和不断改进的过程，随着检验系统的改进而不断变化。

早在1988年，在美国联邦调查局生物实验室主持下由31名美国、加拿大法庭DNA科学家及其它科学家组成了DNA分析方法技术工作组（Technical Working Group on DNA Analysis Methods，TWGDAM），提出了实施质量保证和质量控制体系的一些原则。在以后的几年里不断完善、修改。

欧洲在1989年成立了由欧洲国家警察组织、大学研究所等承担法医个体识别鉴定工作的实验室所组成的欧洲DNA分型小组（EDNAP，European DNA Profiling group），对DNA分型的技术操作及结果判读等提出了质量保证和质量控制体系，以达到分型结果的可比性。1998年，DNA分型在第27届欧洲地区会议中通过新的质量保证和质量控制体系，着重在标准化方面作了详细论述：包括DNA技术，线粒体DNA群体遗传问题，确定了欧洲标准的STR基因座（ESS）vWA、FA31/A、TH01（TC11）、D21S11、FGA；明确了法庭DNA实验室对各类不同检材收集与提取的工作要求，欧洲各国DNA技术的协调，DNA数据库、质量保证与培训和国际间DNA分型交流的必备条件等。

拉丁美洲（墨西哥及其南部的美洲）的20多个国家在1997年共同成立了一个拉丁美洲DNA分型工作组（GITAD），其目的是在所有拉丁美洲的DNA分型实验室起一个协调作用，以便联合参与技术培训和经验交流，促进质量保证（QA）和质量控制（QC）系统。

质量保证及质量控制体系包括了DNA检验的各个环节，既有管理与技术方面，又有人员和检测程序方面，主要包括以下项目：实验室管理、人员素质及培训、仪器设备、试剂材料、证据保存程序、分析方法与程序、水平测试、结果解释、报告的书写和复核、安全措施与审核等。使每个DNA的实验室所有的工作环节都处于受控状态，以确保DNA检验分型的正确性、可比性和实验室数据的完整性。

23.1.2 质量保证体系具体内容

1. 人员

要求分析人员具备一定的学历，了解、熟悉DNA检验相关基础学科如生物化学、遗

传学、分子生物学、群体遗传学及统计学等，经过上岗前培训，具有一定法医 DNA 分析经验，通过资格测试。在工作中接受再教育，如阅读近期的科学文章，参加研讨会、专业会议等了解新知识、新技术，不断更新，提高技术水平。

2. 实验室文件资料

实验室必须保存 DNA 分析过程涉及的各方面的文件以及利于分析、结果解释有关文件与实验记录，作为科学检查、数据重新评估、DNA 程序检查的档案。需要的文件有：

(1) 检验方法和 DNA 分型程序

(2) 群体数据——包括数目、来源及民族

(3) 试剂：记录试剂的批号、厂家和有效期

(4) 案件记录：每个案件检验的原始资料都应记录保存

(5) 数据分析及报告

(6) 物证处理方法

(7) 仪器的使用说明书及维护、校准数据

(8) 资格测试记录、个人培训记录

(9) 方法有效性记录

(10) 质量保证及核查记录

(11) 质量保证手册

(12) 安全守则

3. 实验室要求

(1) 实验室实行严格的物理空间上的隔离。检材的前处理、DNA 的提取、PCR 扩增和产物的分析应当在不同空间内进行，避免检材 DNA 的交叉污染以及外源 DNA 的污染。具体分区为：①样品初检区：检材查验、照相和镜检区域。② 提取区：DNA 提取、浓缩和酶解区域。DNA 含量极低的检材（毛干、陈旧骨骼等）要与其它检材分区或在不同时间提取。③PCR 加样区：PCR 试剂、模板等在此加样。④ 扩增及检测区：加样完毕并盖上盖的试管放人此区的扩增仪中扩增。扩增产物的保存及检测也在此区。此区的仪器和试剂不得挪往其它区域。

(2) 实验室每年按标准检查一次。

(3) 每二年有其他机构参与年检一次。

(4) 实验室需制订且执行书面的诠释实验数据的手册。

4. 仪器校准、维护及设施

(1) 实验室选择适合所开展检验项目的仪器，并处于良好状态，有专人进行仪器维护。

(2) 实验室应备所有仪器的详细使用说明书（包括仪器的生产厂家、型号、到货时间等）及使用记录以备查询。

(3) 定期检修和校准，以保证其正常运转。保存仪器设备的维修记录。

(4) 新购置或维修过的仪器在用于案件分析检验前，应进行校准。

5. 试剂：

(1) 实验室应使用与所用方法相适应的试剂，符合 DNA 分型的技术标准。

(2) 实验室应有明确的试剂配方。配置的试剂应标明：试剂名称、成分、浓度、配制时间、有效时间及配制人等。

(3) 新试剂应经过试验证明结果无误后方可使用。

(4) 实验室在使用以下关键试剂前应确认其最佳量。

①限制性内切酶

②用于 DNA 分型的商品试剂盒

③标准样品

④分子量标准

⑤引物

⑥耐热性 DNA 聚合酶

6. 检验方法

(1) 法医 DNA 实验室应采用经鉴定并被同行业专家所认可的检验方法用于检案。必须经过严格检验后方能认可，包括实验室之间一致性检验、人群调查、环境耐受检验、基质检验和灵敏度检验等有效性确认。

(2) 实验室应有全部检验项目的详细操作手册，其中包括试剂配制、设备、对照、样品处理、DNA 提取、DNA 检测标准程序及数据处理及结果解释。

(3) 有效性确认包括：

①同一性：使用新鲜或一定条件下保存的组织和体液评估分型程序，同一个体不同组织的分型应一致。

②重复性：用已知分型的样品评估技术在实验室内或不同实验室间重复性。

③人群调查：调查不同人种间等位基因分布数据。

④一致性：验证斑迹样品与相应的液体样品的分型一致性。

⑤混合样品分型：检测系统对混合样品分析能力，确定系统的最低量。

⑥环境耐受检验：已知样品暴露于不同环境，评估系统对时间和降解环境（温度、湿度、紫外线）等影响因素的反应。

⑦基质检验：将体液与常见物质（如染料、土壤）混合或涂于织物、毛皮等常见载体上进行检验。

⑧非人样品检验：验证该方法是否能对人以外的生物样品 DNA 进行分型。

⑨灵敏度检验：确定能得到可靠分型所需的最小样本量。

⑩实用性检验：将新方法应用到实际案件中，检验实用性。

⑪遗传性：家系调查证明法医检验中的遗传标记的遗传方式。用于父权鉴定的基因座的突变/重组概率应较低。

⑬基因定位：基因座的染色体定位准确。

⑭检验记录：检测多态性基因座方法的主要点应记录在案。对于 RFLP 基因座应准确描述所用的限制性内切酶及探针；对于 PCR 基因座应准确描述其引物（如果需用探针也应加以描述）。

RFLP 分析程序的有效性确认

①限制性酶切：保证彻底和特异的内切酶消化

②凝胶电泳：建立可重复分离 DNA 片段的参数

③转移/杂交：建立重复转移 DNA 片段的参数，确定一定特异性的杂交和洗脱条件。

④确定片段大小精确测定程序

PCR 及相关检验方法的认可

①扩增 PCR 必须标明引物序列

②扩增样品必须严格保护不被扩增产物污染，扩增反应条件及试剂（引物、酶和离子）浓度应严格控制，保证结果一致性。

③评估和指出潜在的差异扩增，评估复合扩增体系每个等位基因扩增效率。

④PCR 产物检测：每组样品都应设阳性和阴性对照。PCR 产物直接分型，应使用合适的分子量标准标识等位基因。检测时待测样本要在等位基因混合物或分子量标准之间的泳道上电泳；直接测序分析 PCR 产物，要求建立评估序列的适当标准；通过杂交分型，杂交和洗脱条件必须严格统一，保证结果的特异性。

⑤如果 DNA 直接固定到膜上，要用阳性对照检验其是否连接。如果探针直接固定到膜上，必须使用一定的方法确定待测样品中是否含有扩增产物。

7. 物证检材管理

(1) 实验室应有物证检材管理文件，严格遵守，确保物证检材被检测过程的完整性。

①物证应标记和编号，以免混淆。记录或描述物证剪取部位。

②所有物证都有登记表，记录各检测步骤和环节。

③有固定专门的物证检材贮藏空间。

(2) 实验室应保留部分物证检材并妥善保管防止腐败和丢失，以便复核。

8. 分析程序

①进行样品评估与准备。

②DNA 提取，转移样品或加样时要签名或经他人核对。

③DNA 定量分析。

④DNA 分型。实验中设置阴性、阳性及试剂空白对照，必须严防污染。采用等位基因标准物或和内标进行分型。

9. 检验报告

书面记录实验过程与分析结果，用于检验报告书写。

(1) 实验室应保存物证检验记录、原始数据及所有人员的分析结果。

(2) 检验报告应包括：

①送检单位及送检人

②检材描述

③检验方法的描述：所使用的检验方法及仪器、检验项目及试剂

④检验结果，根据人群数据计算匹配概率或相关几率，得出结论

⑤对结论的说明

⑥出具日期

⑦检材处理

⑧检验人员的职称及签名

（3）实验室要书面记录出具检验报告的过程。

10．资格测试

DNA鉴定人员和或实验室应每年接受一次外单位的公开或非公开的鉴定资格测试。检查实验室或鉴定人员是否具有鉴定案件能力。

11．审核

（1）实验室对所有检验文件及报告均需进行行政及技术方面的审核，以保证结论及有关数据的合理性和科学性。实验室应建立一种机制以解决检验人与审核人之间对检验结果持不同结论时的矛盾。

（2）实验室应记录每个检验人员的年度案件鉴定监测结果。

12．错误分型时的处理措施

（1）管理问题：管理问题（样品分型错误、保存不当等）造成明显的错误分型时，应对照质量控制各项规程核查或完善质量控制规程。

（2）系统问题：系统问题（仪器、试剂、实验环境等）造成明显的错误分型时，应立即核查上次合格的测试与本次测试之间检验的所有案件，找到错误原因并纠正后应告知所有实验室人员以免类似错误重复发生。

（3）分析检验问题：分析人员本人的原因造成明显的错误分型时，该检验人员应停止检案，DNA实验室应对该检验人员上次合格的测试与本次测试之间检验的案件适当核查。找到问题所在并纠正后，该检验员经再次的资格测试合格后方可检案。

13．检查监督

从事DNA检验的实验室应遵循质量保证体系，确保数据的正确性、完整性和可信性。

23.2 DNA数据库

23.2.1 数据库的意义

样品与嫌疑人直接比对只发挥了DNA分型的极小部分作用，DNA数据库可发挥DNA的较大作用。DNA数据库将法庭科学与计算机技术联合起来，DNA分型结果输人计算机网络系统构建的DNA数据库，通过网络系统，将DNA数据连接起来，实现数据共享，通过比对系统进行比对，确定数据库是否存在与比对样品分型一致的个体，如果与库中某一嫌疑人匹配时，就可缩小排查范围，甚至直接认定案犯；也可以将几起案件串并起来，迅速而有效地为侦查破案服务，发挥DNA在破案上的更大作用。荧光标记STR复合扩增技术及全自动DNA分型仪器的出现，使STR－PCR技术检测结果的自动化、数码化，加速了大规模建立DNA数据库的进程。

采集暴力犯罪人员的DNA样本，有犯罪前科人员的DNA样本进行分析构建的DNA数据库，简称前科库。从犯罪现场提取的生物物证进行DNA分析构建的DNA数据库，称为现场DNA数据库。现场检材DNA分型通过与现场库中的数据进行比对，两者匹配时可以进行串并案件，为侦破多次做案的系列案件提供了科学证据与侦破线索。与前科库中所有样本进行比对，当现场检材各基因座基因型与库中某一嫌疑人匹配时，再次分析库中被比对中的样品DNA分型，确认所有分型结果正确无误，可以直接认定案犯。此外，建立失踪人员父母DNA数据库，将无名尸、碎尸DNA检验数据与失踪人员父母DNA数据库中数

据进行比对，可进行尸源认定。将被拐卖或失踪儿童 DNA 检验数据与失踪儿童父母 DNA 数据库中数据进行比对，可寻找并认定他们的亲生父母。

DNA 数据库技术具有巨大的发展潜力及广阔的应用前景，而且随着数据库的科学管理和标本数据的增加，数据库的有效利用率会越来越高，其在侦察破案方面发挥的作用将会递增。建立犯罪 DNA 数据库已成为法医物证检验技术最主要的发展方向之一。如何建立及使用本地区、本国的 DNA 数据库已引起各个国家及政府相关职能部门的高度重视。

23.2.2 国外 DNA 数据库概况

英国是世界上 DNA 建库最早的国家，于 1990 年开始正式建立 DNA 数据库，主要对犯人及未破案件现场生物检材建库，1993 年英国皇家刑事审判专门调查委员会建议警察可以提取嫌疑人非体内样本用于 DNA 检验。1994 年《刑事审判与公共秩序法》授权 FSS 进行 DNA 数据库的研究，1995 年 4 月该条例正式成为英国法律，已成功建立了 35 万人的 DNA 库。

美国联邦调查局（Federal Bureau of Investigation，FBI）于 1990 年联合 8 个州 12 个实验室建立 DNA 联合索引系统（Combined DNA Index System，CODIS）。CODIS 系统是由美国联邦调查局建立和管理的 DNA 数据库，使地方、州和联邦实验室之间能够以网络方式来交换、比较 DNA 分型，从而达到罪犯信息异地查询的目的。CODIS 索引系统通过两个索引来发挥这种导向作用：罪犯 DNA 索引（Convicted offender DNA Index）和法庭索引（Forensic Index）。CODIS 是一个庞大的系统，组织形式上分为国家、州、地方三级管理，以金字塔式结构构成。地方 DNA 分析系统（Local DNA analysis ststem，LDAS）为金字塔的基底，其作用是管理样本，输入 PCR 分型（PCR typing）结果，将数据核查无误后传送给地方 DNA 索引系统（Local DNA Index System ，LDIS），建立地方的数据库。州 DNA 索引系统（State DNA Index System，SDIS），为金字塔的中间结构，它从各地方 DNA 索引系统中挑选数据上载而建立的 DNA 数据库。它也有两个索引：州案件索引和州罪犯索引。国家 DNA 索引系统（National DNA Index System，NDIS）是金字塔的顶尖，同样，它是由州 DNA 索引系统中挑选的数据上载而建立的数据库，它也包括国家案件索引（收集地方的案件数据）和国家罪犯索引。国家 DNA 索引系统（National DNA Index System）由 FBI 直接管理运作，并负责 CODIS 系统软件研究、安装、调试和培训等。1998 年 CODIS 并入美国法庭科学实验室互联网（LABNET）。地方 DNA 索引系统设立在当地警察局，DNA 检验人员将检验结果通过 CODIS 软件进行分析，并与地方实验室检验的有关结果进行匹配检验，若将数据发至州 DNA 数据库，地方实验室管理者可与州内参与 CODIS 的其他实验室实现数据共享。州 DNA 索引系统可实现州内各实验室内数据共享与检索查询，将数据发送到国家数据库后，则可与其他州实现数据共享。1991 年美国国会通过了关于建立 DNA 数据库的法案，至 1996 年已有 40 个州 46 个实验室参与了 CODIS，98 年 50 个州 94 个实验室加入 CODIS，至 2000 年有 50 个州 130 个实验室加入 CODIS 系统。

日本情况与美国相似，已初步建立“犯罪 DNA 数据库”。

23.2.3 建库内容与要求

建立 DNA 数据库的内容：分析方法简单、易自动化和遗传信息量大的遗传标记的确定，较易实现标准化分型方法的确定，法律程序化样品的采集，正确无误的数据录入与传

送，数据搜索与比对系统确定，数据库比对结果的应用与解释。

标准化是建立 DNA 数据库的前提及基础，是建库的首要环节。DNA 数据库能否成功运行的关键是不同实验室、不同实验者对同一样本测试得到的基因型完全一致，具有可比性，否则就不可能为法庭科学提供正确的信息及线索。只有实行标准化及质量控制，才能建立高质量的数据库。如果进入数据库的数据有误，有可能导致假排除或假认定两类错误，使罪犯逍遥法外或造成冤假错案。所以对技术方法的筛选、遗传标记的确定、参数指标制定、人员及实验条件的审核、质量保证体系的建立等都是围绕数据正确这一中心问题。

输入数据正确无误作为首要工作，建立可操作的严格的流程模式并将其制度化，在人员方面对人员进行专门培训及再训练，并实行严格的淘汰性测试，在样本采集方面设计专门的工具盒及个人信息记录表格，制作采样培训录像带，将所有个人信息由录入人员录成一个电子表格，然后输入到数据库中并与 DNA 分型结果相关联。对样本存放、DNA 提取方法、定量、扩增 DNA 模板用量、扩增程序、试剂、电泳分析方法、设备等均做详尽的规定。为保证等位基因型判读的准确性，等位基因分析由 2 名操作员独立进行。每个月取检测样本的 10%用于盲测，每周还将不完全匹配的样本进行多至 13 ~ 14 个基因座的检测。

主要参考文献

1. 郭景元，李伯龄主编. 中国刑事科学技术大全·法医物证学. 北京：中国人民公安大学出版社，2002

2. John M. Butler. Forensic DNA typing. Academic press, A Harcourt Science and Techonology Company, London, UK, 2001

3. Epplen J, Lubjuhn T. DNA profiling and DNA fingerprinting. Birkhauser Verlag Basel/Switzerland, 1999, 53～68

4. 吴梅筠主编. 法医物证学. 北京：人民卫生出版社，1998

5. Patrick J. Lincoln and Jim Thomson: Forensic DNA profiling protocols. 1998, Human Press, Totowa, New Jersey

6. 郭景元主编. 法医物证学. 北京：中国人民公安大学出版社，1997

7. Keith Inman, Norah Rudin. An Introduction to Forensic DNA Analysis, CRC Press Boca Raton, New York, 1997

8. 杨庆恩. DNA 在法庭科学中的应用. 北京：中国人民公安大学出版社，1994

9. Brown TA. DNA Sequencing. Oxford University Press, New York, 1994

10. Lorne T. Kirby. DNA Fingerprinting - An Introduction. Stockton Press, 1990

11. 贺林主编. 解码生命. 北京：科学出版社，2000

12. 马立人，蒋中华主编. 生物芯片. 北京：化学工业出版社，1999

13. 迪芬巴赫 CW，德维克斯勒 GS. PCR 技术实验原理. 科学出版社，1999

14. 朱玉贤，李毅编著. 现代分子生物学. 北京：高等教育出版社，1997

15. 林万明. PCR 技术操作和应用指南. 北京：人民军医出版社，1993

16. 金冬雁，黎孟枫等译. 分子克隆—实验指南，第二版. 北京：科学出版社，1992

17. 王美岭，韩金祥编著. 基因合成极其应用. 山东大学出版社参考文献，1991

18. Sambrook J, Fritsch EF and Maniatis T. Molecular Cloning —a laboratory manual, 2nd, 1989; Cold Spring Harhor Laboratory Press, New York

19. 中山大学编. 生化技术导论. 北京：人民教育出版社，1981

20. 张龙翔等. 生化实用方法和技术. 北京：人民教育出版社，1981

21. 沈同，王经岩，赵邦娣主编. 生物化学. 北京：人民教育出版社，1980

22. 王海云，冯北元译. 实用生物化学原理和技术，北京，科学出版社，1979

23. 李春喜，王志和，毛文林. 生物统计学. 北京：科学出版社，2000

24. 陈国弟，侯一平，李英碧等. 人类 D19S400 基因座在不同人种中的遗传多态性研究. 法医学杂志，1999，15：80～81

25. 陈红，姜成涛，刘键. 宁夏地区回族三个复合 STR（D16S539，D7S820，D13S317）扩增多态性的研究. 第二次全国法医物证学学术交流会论文集，1998，105～107

26. 陈慧，余纯应，杨容芝等. 武汉地区汉族人群 CSF1PO 座位多态性分析. 法医学杂志，1999，15：84～85

27. 陈松，胡兰，季安全等. COL2A1 和 IL-6 两个位点扩增片段长度多态性法医学应

用的研究．第五次全国法医学术交流会论文集，1996，285～287

28. 陈松，李伯龄，倪锦堂等．中国人P33.6位点扩增片段长度多态性．中国法医学杂志，1993，8（3）：129～131

29. 程保文，刘玲，苏少明等．云南汉族p33.6位点的基因频率调查．中国法医学杂志，1996，11（4）：217—218

30. 程保文，邹浪萍．HUMCYAR04基因座扩增片段长度多态性．中国法医学杂志1997，12（3）：161～163

31. 程俊峰，郑秀芬，叶健等．广西壮族人群2个Y－STR基因座的多态性分析及其法医学应用．中国法医学杂志，2001，16（3）：166～167

32. 郭大玮，郎海丽，徐晓莉．人类vWF基因内含子40（HUMFA31（C），nt31/2215－2380）STR多态性结构以及在中国和后来荷兰群体中的分布．第二次全国法医物证学学术交流会论文集，1998，89～94

33. 韩宇，李生斌．国人D17S30遗传多态性研究．第五次全国法医学术交流会论文集.1996，275～276

34. 侯一平，李英碧，吴谨等．30个STR遗传标记的法医遗传学可靠性评估。见：侯一平，刘世沧．法医学进展与实践（第二卷）．第一版．成都：成都科技大学出版社，1999

35. 侯一平，刘雅诚，李剑波等．法医DNA分型的可重复性研究．中国法医学杂志1998，13：163～165

36. 侯一平，吴谨，李英碧等．Y染色体特异短串联重复序列初步研究．中华医学遗传杂志，1999，16：65～69

37. 侯一平．中国人Col2A1基因座的扩增片段长度多态性．遗传学报，1995，22（4）：245～251

38. 胡兰，倪锦堂，陈松等．PYNZ22位点扩增片段长度多态性法医应用的研究．中国法医学杂志，1995，10（1）：5～7

39. 霍振义，刘雅诚，唐晖等．三个STR位点复合扩增多态性及法医学应用．中国法医学杂志，1998，13：39～41

40. 季安全，李伯龄，倪锦堂等．P33.4位点扩增片段长度多态性及其法医学应用的研究．中国法医学杂志，1994，9（4）：210～214

41. 季安全，李继周，叶健等．JL－02多位点探针DNA指纹的法医学应用研究．中国法医学杂志，1997，12（2）：65～68

42. 季安全，赵鑫，李继周等．Profiler Plus系统在法医学DNA检验中的问题探讨．首届全国法医DNA检验技术研讨会，2000：83～87

43. 季安全，叶健，郑秀芬等．D17A1探针点杂交DNA定量．刑事技术，2001，1：10～11

44. 季安全，叶健，郑秀芬等．荧光标记STR复合扩增系统的研究．刑事技术，2001，12～16

45. 江斌，郭景元，梁贤猷．运用变性PAGE和荧光检测技术对广州汉族人群六个

STR 位点的研究．法医学杂志，1999，15（3）：141～143

46. 姜先华，吕世惠，王国林等．DNA 指纹图在法医学鉴定中应用的研究（Ⅰ）．中国法医学杂志 1990，5（1）：11～14
47. 姜先华，王国林，韩景松等．根据 y－染色体特异性 DNA 序列鉴定牙齿性别．中国法医学杂志，1990，5（3）：133～135
48. 进藤祥子，郑秀芬，大谷真纪等．日本人のトリプトオファンヒドロキシラ一ゼ遺传子の多型．DNA 多型，7 卷，东洋书店．1999：223～226
49. 景强，孙光云，吴梅筠．成都地区汉族群体 ApoB 3′VNTR 位点长度多态性研究．中国法医学杂志，1995，6（1）：1～2
50. 厥庭志，李莉，林源等．用多重 PCR 及荧光检测技术分析 CTT 位点的多态性．法医学杂志，1999，15（3）：144～146
51. 蓝翎，张小为，瞿振义等．用寡核苷酸探针（CAC）5/（GTG）5 进行人的 DNA 指纹分析．遗传学报，1993，20（1）：1～7
52. 李伯龄，丁焰，倪锦堂等．ApoB 位点扩增片段长度多态性及其应用．中国法医学杂志，1992，7（1）：1
53. 李伯龄，丁焰，倪锦堂等．pMCT118 位点扩增片段长度多态性及其应用．中国法医学杂志，1991，6（3）：132～135
54. 李伯龄，倪锦堂，储小兰等．DNA 探针在法医学中的应用研究—用 PYNH24 探针作个体识别和家系分析．遗传学报，1989，16（6）：470～474
55. 李伯龄，倪锦堂，储小兰等．DNA 纹印技术的初步研究．中国法医学杂志 1988，3（1）：4～7
56. 李伯龄，倪锦堂，储小兰等．用 α－珠蛋白－3′HVR 探针研究 DNA 指纹．中国法医学杂志，1989，4（1）：5～7
57. 李路平，高浔，陈少恒等．青岛地区人群 CSF1PO－TPOX－TH01 基因座的基因频率调查．中国法医学杂志，1999，14：102
58. 李万水，胡兰，陈松．煮熟组织块的种属鉴定 1 例．中国法医学杂志，1999，14（1）：47～48
59. 李晓平，杜庆新，顾明波等．ABO 基因座限制性扩增片段长度多态性的研究．中国法医学杂志，1998，13（4）：197～201
60. 李晓平，杜庆新，顾明波等．ABO 基因座限制性扩增片段长度多态性在法科学中的应用．中国法医学杂志，1999，（1）：30～32
61. 廖林川等．用温度调控高效液相色谱探索基因组单核苷酸多态性的方法研究．中华医学遗传杂志，2000，17（3）：204～207
62. 刘超，王穗保，梁赏猷等．陈旧骨骼 DNA 提取及性别鉴定．法医学杂志，1998，14（4）：195～200
63. 刘健，尤朝阳，郑秀芬等．复合扩增 STR 位点片段长度多态性的研究．中国法医学杂志，1994，9：223～225
64. 刘明俊，巩岩，张少华等．国人六个高变 DNA 位点的基因频率分布．中国法医学

杂志，1994，9（4）：219～221
65. 刘明俊．200例国人DNA纹印基因频率分布—D2S44及D17S79位点的DNA多态性．西安医科大学学报，1989，10（2）：118
66. 刘万清，贺林．遗传作图的第三代工具——DNA芯片．生命科学，1998，10（2）103～105
67. 刘雅诚，霍振义，唐晖等．北京汉族群体9个STR位点的频率分布及其法医学应用．中国法医学杂志，1998，13（4）：210～212
68. 吕嵩，黄晓琳．DNA芯片技术原理及在遗传学研究中的应用．国外医学遗传学分册，1998，21（6）284～289
69. 孟海英，侯一平，陈国弟等．D8S384基因座的遗传多态性及其在法医物证检验中的应用．中国法医学杂志，1999，14：129～134
70. 倪锦堂，姚翔，郭红玲等．PMLJ14～HaeⅢ单位点DNA图谱技术在法医学中的应用．中国法医学杂志，1993，8（1）：3～5
71. 倪锦堂，姚翔，郭红玲等．人体基因组D14S13位点限制性片段长度多态性的研究．遗传学报，1993，20（2）：97～101
72. 倪锦堂，姚翔，郭红玲等．三个单位点DNA杂交联合应用的研究．中国法医学杂志，1994，9（4）：215～217
73. 倪锦堂，叶健，刘健等．非同位素标记DNA探针的初步研究．中国法医学杂志，1991，6（3）：129～130
74. 倪锦堂，叶健，刘健等．辣根过氧化物酶标探针及光增强法检测DNA指纹图的法医学研究．遗传学报，1992，19（3）：193～197
75. 平原，顾丽华，陈连康．STR系统HUMCSF1PO～HUMTPOX～HUMTH01三位点复合扩增在上海地区的频率分布．中国法医学杂志，1998，13：37～38
76. 邱宁，李万水，陈松等．ABO基因型检验及其法医学应用研究．法律与医学杂志，1999，6（1）：29～30
77. 孙开来，吴东宁．DNA芯片的发展和意义．国外医学遗传学分册1998，21（6）：281～284
78. 唐晖，刘雅诚，朱晓君等．33.15及33.6探针对中国人DNA指纹图的研究．法律与医学杂志，1998，5（3）：109
79. 唐双柏，郭景元，刘超等．广州地区汉族人群Y染色体多个STR位点的研究．法医学杂志，1999，15（2）：86～88
80. 吴梅筠，孙光云．分析扩增DNA序列测定人发的性别．法医学杂志，1991，7（1）：1～3
81. 吴梅筠，孙光云．体外DNA扩增技术鉴定血痕性别三例报告．中国法医学杂志，1991，6（2），68～69
82. 吴微微，程建波，郑秀芬．浙江省汉人D1S80位点扩增片段长度多态性群体调查．中国法医学杂志，1998，13（4）：233～234
83. 吴微微，郑秀芬等．人类D1S80位点5’旁侧Hinf I酶切位点．中国刑警学院学

报，1998，4：44～45

84. 吴微微，郑秀芬．一种改良 Chelex 方法提取血痕 DNA．刑事技术，1999，(1)：20～21
85. 伍新尧，杨英浩，罗超权．Y 染色体特异性探针作斑点印迹检测人血痕性别．法医学杂志，1991，2：7～8
86. 肖仰哲，王保捷，刘利民等．vWF Ⅲ基因座扩增片段长度多态型及其次法医学应用技术．中国法医学杂志，1999，14 (1)：23～25
87. 徐振波，吴梅筠，孙光云等．中国成都地区汉族群体四核苷酸重复序列 FIBR 基因座的多态度性．中国法医学杂志，1999，Vol.14 (1)：19～23
88. 阎威，贾静涛，姜先华等．短串联重复位点 ACTBP2 (SE33) 的扩增片段长度多态性研究．中国法医学杂志，1996，11：129～133
89. 杨庆恩等．复合 PCR－RFLP 技术检测 ABO 基因型．中华医学遗传学杂志．1999，(2)，110～112
90. 杨庆恩．亲子鉴定 DNA 分型亲权指数的简化计算法．中国法医学杂志，1998，13 (2)：90
91. 叶健，丁焰，倪锦堂等．混合斑分离的应用研究．刑事技术，1991；(3)：1
92. 叶健，季安全，郑秀芬等．陈旧性骨骼 DNA 提取技术的研究．中国刑事科学技术协会第三届学术研讨会论文汇编．北京：中国人民公安大学出版社，2001
93. 叶健，张纯斌，胡兰等．荧光测序分型技术检验陈旧骨骼和牙齿的 STR 位点及 HumamLPL 位点的基因频率调查．法医学杂志，2000，16 (1)：16～17
94. 叶健，郑秀芬，陈松等．DNA 技术在法医物证鉴定中应用的经验．中国法医学杂志，1996，11：37～40
95. 叶健，郑秀芬，倪锦堂等．HinfⅠ和 HaeⅢ酶解条件对 DNA 指纹图的影响．中国法医学杂志 1995，10 (4)：218
96. 叶健，郑秀芬，周静等．PCR 常见技术的常见问题与处理方法．刑事技术，1999，6：35～36
97. 叶健，郑秀芬，季安全等．熔解曲线分析插入、缺失及 SNP 多态性．中国法医学杂志，2001，16 (4)：214～217
98. 余纯应，杨庆恩，杨容芝等．STR 位点 D6S366 的多态性分析．法医学杂志，1998，14：12～13
99. 余纯应，杨庆恩，朱传红等．汉族人群五个 STR 位点基因座的多态性调查．中国法医学杂志，1997，12：69～73
100. 张纯斌．中国人在 Profiler Plus 复合扩增检验中出现的罕见等位基因．首届全国法医 DNA 检验技术研讨会，2000：69～71
101. 张纯斌，周静，宋慧娟等．用 PCR－测序法对线粒体 DNA 多态区的研究．中国法医学杂志，1995，10 (3)：129～132
102. 赵兴春，姜成涛，郑秀芬等．线粒体 DNA 分型技术．刑事技术，2001，6：29～32

103. 郑秀芬，王小娅 .Y－染色体及其在法医鉴定中的应用 . 中国法医学杂志，1998，13 (3)：183～187

104. 郑秀芬，吴微微，谢志萍等. 中国汉族群体 Y 染色体 DY390 位点基因频率调查 . 法医学杂志，1999，15 (1)：33～34

105. 郑秀芬，叶健，倪锦堂等. D1S8 位点 MVR～PCR 分型及其法医学应用 . 中国刑事科学技术协会首届学术研讨会论文汇编，北京：群众出版社，1996

106. 郑秀芬，黄正光，纪贵金等. DNA 芯片及在法医物证鉴定中的应用展望 . 中国法医学杂志，2000，15 (2)：112～114

107. 郑秀芬，叶健，倪锦堂等. HPR 标记探针 MVR 分型法在法医学上的应用 . 法医学杂志，1998，14 (1)：14～15

108. 郑秀芬，程建波，徐林苗等. 扩大法医 DNA 分析——其他法医 DNA 分析 . 中国刑事科学技术协会第三届学术研讨会论文汇编 . 北京：中国人民公安大学出版社，2001

109. 郑秀芬，程俊峰，叶健等. Y－STR A10、A7.1、A7.2 和 C4 基因座在法医学应用的研究 . 中国刑事科学技术协会第三届学术研讨会论文汇编 . 北京：中国人民公安大学出版社，2001

110. 郑秀芬，吴微微，徐林苗等. 混合班染色涂片的 DNA 分型 . 中国法医学杂志，1998，14 (3)：235～236

111. 郑秀芬，吉冈尚文，徐林苗. 头皮屑 DNA 分型 . 中国法医学杂志，2000，15 (2)：92～93

112. 郑秀芬，程建波，吴微微等. 已显现指纹的性别鉴定 . 中国刑事科学技术协会第二届学术研讨会论文汇编，北京：中国人民公安大学出版社，1998

113. 郑秀芬，叶健，倪锦堂等. 用 DNA 纯化试剂分离除去 PCR 抑制剂 . 刑事技术，1999，5：16～17

114. 郑秀芬，吉冈尚文 .The polymorphism of CA repeats in Japanese and Chinese populations. Acta Crim.Japan，2000，66 (2)：41～45

115. 郑秀芬，季安全，程俊峰等. Y 染色体 A10－STR 基因座在中国汉族群体的多态性 . 刑事技术，2001，5：27～28

116. 郑秀芬，纪贵金等. 二组分混合样品 DNA 结果解释，中国法医学杂志，2000，15 (4)：203～207

117. 郑秀芬，姜成涛，赵兴春 .STR 自动荧光分析法用于强奸案中混合血痕一例 . 法医学杂志 2000，16 (3)：180

118. 郑秀芬，刘寰，程建波 . 已显现指纹 DNA 分析 . 刑事技术，1998 (1) 1～3

119. 郑秀芬，王琳 . 线粒体 (CA) n 重复子的多态性——中国汉人的群体数据 . 中国法医学杂志，1999，14 (3)：159～160

120. 郑秀芬，吴微微，叶健等. 中国汉族群体 Y 染色体 DY390 位点基因频率调查 . 法医学杂志，1999，33 (1)

121. 郑秀芬，吴微微，纪贵金等. Y 染色体 C4 STR 基因座在中国汉族群体中的多态

性．中国法医学最新科研与实践—全国第六次法医学术交流会论文精选．北京：中国人民公安大学出版社，2000

122. 郑秀芬，徐林苗．STR 位点的法庭科学应用现状．公安大学学报，1999（2）
123. 郑秀芬，叶健，委安全等．PEP－PCR 法及法医学应用可行性研究．中国法医学杂志，2000，15（1）：18～21
124. 郑秀芬，叶健，倪锦堂等．数字编码小卫星可变重复序列法对中国汉人分型，中国法医学杂志，1996，11（2）：72～75
125. 郑秀芬，程俊峰，叶健等．汉族和壮族人群在 4 个 Y－STR 基因座的遗传多态性研究．遗传学报，2002，2：95～100
126. 郑秀芬，叶健，周静等．荧光标记检测 STR 复合基因座的应用．刑事技术，2000（1）：25～28
127. 郑秀芬，赵兴春，姜成涛等．9 个 STR 基因座在中国汉族群体中的多态性及其法医学应用．中国法医学杂志，2001，3：53～57
128. 周静，张纯斌，倪锦堂．用 PCR 扩增 X－Y 特异性片段性别鉴定的研究．中国法医学杂志，1993，8（4）：201～203
129. 朱少建，等．广西壮、汉人群 9 个 STR 基因座的遗传多态性及其法医学应用．首届全国 DNA 检验技术研讨会论文集．2000
130. 朱少建，刘超，李发贵等．壮族人群 3 个 STR 基因座基因频率分布及其法医学应用．中国法医学杂志，1999，14：15～18
131. 邹浪萍，李正建，郑秀芬等．增强荧光标记 33.6 探针进行 DNA 指纹分析．中国刑警学院学报，1997，57（3）：31～33
132. Aspollu A, Kelve M. The first criminal case in estonia with dog' s DNA dataadmitted as evidence. 19th International congress of international society for forensic genetics, 2001 Munster, Germany
133. Albarran, C.O.Garcia, R. Deka. Analysis of D1S80 VNTR allele polymorphism and association with a nearby flanking sequence polymorphism in two Spanish population. Advance in Forensic Haemogenetic, 6, Springer～Vert/lag, Berlin, 1996: 151～153
134. Andersen, JF et al. Report on the third EDNAP collaborative STR exercise. Forensic Sci. Int., 1996, 78 : 83 ～93
135. Antonio Alonso, Pablo Martin, Cristina Albarran et al. A Hinf I polymorphism in the 5′ flanking regioon of human VNTR locus D1S80. Int J Leg Med, 1995, 107: 216～218
136. Balazs I et al. Human population genetic study of five hypervariable DNAloci. Am J Hum Genet, 1989, 44: 182
137. Bar W et al. DNA Commission of the International Society for Forensic Genetics: Guideline for mitochondrial DNA typing. Int J Legal Med, 2000, 113: 193～196
138. Beecher Jody E, McGall Glem H et al. Chemically amplified photolithography for the fabrication of high density oligonucleotide assays. Polym Mater Sci Eng, 1997, 76: 597～598
139. Benecke M. Randon amplified polymorphic DNA（RAPD）typing of necrophageous insects

(dipteramcoleoptera) in criminal forensic studies: validation and use in practice. Forensic Scie Int., 1998, 98: 157 ~ 168

140. Biggin MD, Gibson TJ, Hong GF. Buffer gradient gels and ^{35}S label as an aid to rapid DNA sequence determination. Proceedings of the National Academy of Sciences USA, 198, 80: 3963

141. Boerwinkle E. Rapid typing of tandemly repeated hypervariable loci by the polymerase chain reaction: Application to the apoliprotein B 3′ hypervariable region. Proc Natl Acad Sci USA, 1989, 86: 212

142. Brinkmann C, Forster P, Schurenkamp M. Human Y – chromosomal STR haplotypes in a Kurdish population sample. International Journal of Legal Medicine, 1999, 112 (3): 181 ~ 183

143. Budowle B et al. Fixed – bin analysis for statistical evaluation of continuous distribution of allelic data from VNTR loci for use in forensic comparisons. Am J Hum Genet, 1991, 48: 841

144. Caglia A, Dobosz M, Boschi I. Increased forensic efficiency of a STR – based Y – specific haplotype by addition of the highly polymorphic DYS385 locus. International Journal of Legal Medicine, 1998, 111 (3): 142 ~ 146

145. Chakraborty R et al. Statistical interpretation on DNA typing data. Am J Hum Genet, 1991, 49: 895

146. Chee Mark, Yang Robert. Assessing genetic information with high – density DNA arrays. Science, 1996, 274: 610

147. Csete K, Szabo A, Varga T. The Y – STRs DYS19 and DYS390 in a south – east hungarian (Szeged area) population. International Journal of Legal Medicine, 1999, 112 (3): 207 ~ 208

148. David D. Options available – from start to finish – for obtaining expression data by microarray. Nature genetics supplement, 1999, 21: 25 ~ 32

149. DNA Commission of the ISFH Report concerning further recommendations of the DNA Commission of the ISFH Regarding PCR – based polymorphisms in STR (short tandem repeats). Int J Legal Med, 1994, 107: 159 ~ 160

150. Edwards A, Civitello A, Hammond HA et al: DNA typing and genetic mapping with trimeric and tetrameric tandem repeats. Am J Hum Genet, 1991, 49: 746 ~ 56

151. Eriksen B, Svensmark O. Comparison of two molecular weight markers used in DNA – profiling. Int J Legal Med, 1992, 105 (3): 145 ~ 8

152. Feinberg AP and Vogelstein B. A technique for radiolabeling DNA restriction endonuclease fragment to high specific activity. Anal. Biochem., 1983, 132: 6

153. Francisco LV, LengstonAa, Mekkersh CS et al. A class of highly polymorphic tetranucleotide repeats for canine genetic mapping. Mammalian Genome, 1996, 7: 359 ~ 362

154. Fridez F, Rochat S, Coquoz R. Individuanl identification of cats and dogs using mitochon-

drial DNA tandem repeats? Science & Justice, 1999, 39: 167 ~ 171

155. Gehrig C et al. Allele frequency of 3 chromosome Y - specific STR loci and their application to forensic casework. Proceedings of the 6th international symposium on human identification. Scottsdale. USA, 1995

156. Gene E, Pique N, Borrego A et al. Catalonían population study of thetetranucleotide repeat loci D3S1358, D8S1179, D18S51 and D19S253. Int J Legal Med 1998, 112: 75 ~ 77

157. Gene M, Huguet E, Moreno P, Corbella D J et al. Population study of the STRs HUMCD4 and HUMF13A1 in catalonia (Northeast spain) . Int. J Legal Med, 1997, 110: 44 ~ 45

158. Germer S et al. Single - tube genotyping without oligonucleotide probes. Genome Research, 1990, 9: 72 ~ 78

159. Gill P et al. Report of the European DNA profiling group (EDNAP) ——towards standardisation of short tandem repeat (STR) loci. Forensic Sci. Int., 1994, 65: 51 ~ 59

160. Gill P et al. Automated short tandem repeat (STR) analysis in forensic casework——a strategy for the future. Electrophoresis, 1995, 1 6 : 1543 ~ 1552

161. Gill P. The utility of "substrate controls" in relation to "contamination" . Forensic Sci. Int., 1997, 85: 105 ~ 111

162. Gillian Tully, Kevin M, Sullivan et al. Rapid detection of mitochondial sequence polymorphism using multiplex solid - phase fluorescentminisequencing. Genomics, 1996, 34: 107 ~ 13

163. Gotaro Watanabe, Kazuo Unetsu, Isao Yuasa et al. Nucleotide substitution in the 5′flanking region of D1S80 locus. Int. J Med., 1997, 89: 75 ~ 80

164. Has - Rochholz H, Weiler G. Additional primer sets for an amelogenin gene PCR - based DNA - sex test. Int Legal Med, 1997, 110: 312 ~ 315

165. Hacia JG et al. Determination of ancestral alleles for human single - nucleotide polymorphisms using high - desity oligonucleotide arrays. Nature Gentics, 1999, 22: 164 ~ 167

166. Hammer MF . A recent insertion of an Aluelement on the Y chromosome is a useful marker for human population studies. Mol Biol Evol, 1994, 11 : 749 ~ 761

167. Hammer MF. A recent common ancestry for human Y chromosomes. Nature, 1995, 378: 376 ~ 378

168. Hammond HA et al. Evalution of 13 short tandem repeat loci for use in personal identification application. Am J Hum Genet, 1994, 55: 175 ~ 189

169. Holly A, Hammond, Li Jin, Zhong Y et al. Evaluation of 13 Short tandem repeat loci for Use in personal identification applications. Am. J Hum. Genet, 1994, 55: 175 ~ 189

170. Hoogendoorn B et al. Cheap, accurate and rapid allele frequency estimation of single nucleotide and DHPLC in DNA pools. Hum Genet, 2000, 107: 488 ~ 493

171. Horn GT, Richards B, Klinger KW. Amplification of a highly polymorphic VNTR segment by the polymerase chain reaction, Nucleic Acids Research, 1989, 17 (5): 2140

172. Horsell J, Cordiner SJ, Maas EW et al. Forensic comparison of soils by microbial communi-

ty DNA profiling. Science & Justice, 2001, 41 (3): 236

173. Howell WM et al. Dynamic allele - specific hybridization. Nature biotechnology, 1999, 17: 87 ~ 88

174. Huang N E, Budowle B. Fixed Bin Population Data for the VNTR loci D1S7, D2S44, D4S139, D5S110, and D17S79in Chinese from Taiwan. Journal of Foeensic Science, 1995, 40 (2): 287 ~ 290

175. Innis M, Myambo K, Gelfand et al. DNA sequencing with Thermus aquaticus DNA polymerase and direct sequencing of polymerase chain reaction - amplified DNA. Proceedings of the National Academy of Sciences USA, 1988, 85, 9436

176. Innis M, Myambo K, Gelfand, and Browon M. DNA sequencing with Thermus aquaticus DNA polymerase and direct sequencing of polymerase chain reaction - amplified DNA. Proceedings of the National Academy of Sciences, USA, 1988, 85: 9436

177. Jefferys A J et al. The efficiency of multilocus DNA fingerpint probe for indivualization and establishement of family relationship determined from extensive case worke. Am J Hum Genet, 1991, 48: 824

178. Jefferys A J, Ohn F Y, Brookfield. Positive identifcation of an immigration test case using human DNA fingerprints. Nature, 1985, 317: 818 ~ 819

179. Jefferys AJ et al. DNA 'fingerprints' and segregation analysis of multiple markers in human pedigree. Am. J. Hum Genet, 1986, 39: 11 ~ 424

180. Jefferys AJ et al. Spontaneous mutation rates to new length alleles at tandem - repetitive hypervariable loci in human DNA. Nature, 1988, 332: 278 ~ 281

181. Jefferys AJ, Victoria Wilson and Swee Lay Thein. Hypervariable 'minisatellite' regions in human DNA. Nature, 1985, 314: 67 ~ 73

182. Jeffreys AJ. Amplification of humanminisatellites by PCR, towards DNA fingerprinting of singlecells. Nucleic Acids Research, 1988, 16 (23): 10953

183. Jobling MA et al. Recurrent duplication and deletion polymorphisms on the long arm of the Y chromosome in normal males. Hum Mol Genet, 1996, 5: 1767 ~ 1775

184. Johanna K et al. High ~ throughput SNP detection by using DNA pooling and denturing high performance liquid chromatography (DHPLC). Hum Genet, 2000, 107: 483 ~ 487

185. Junge A, Madea B. Population studies of the Y ~ chromosome specific polymorphisms DYS19, DYS389 I + II, DYS390 and DYS393 in a western german population (Bonn area). Forensic Science International, 1999, 101 (3): 195 ~ 201

186. Junge B, Madea. Validation studies and characterization of variant alleles at the short tandem repeat locus D12S391. Int J legal med, 1998, 112: 67 ~ 69

187. Karafet T et al. Different patterns of variation at the X - and Y ~ linked microsatellite loci DXYS156X and DXYS156Y in human populations. Hum Biol, 1998, 70: 979 ~ 992

188. Kayser M et al. Evaluation of Y chromosomal STRs: a multicenier study. Int J Legal Med, 1997, 110: 125 ~ 133

189. Kimpton C P et al. Report on the second EDNAP collaborative STR exercise. Forensic Sci. Int., 1995, 71: 137 ~ 152

190. Klintschar M , Glock B, Dauber E M et al. Genetic variation and sequence studies of a highly variable short tandem repeat at the D17S976 locus. Int J Legal med, 1998, 112: 50 ~ 54

191. Klintschar M , Neuhuber F. A study on the short tandem repea system ACTB2 (SE33) in an Austrian population sample. Int. J Legal Med., 1998, 111: 46 ~ 48

192. Koskinen MT, Bredbacka P. A convenient and efficient microsatellite - based assay for resolving parentages in dogs. Animal Genetics, 1999, 30, 148 ~ 149

193. Lareu MV, Barral S, Salas A et al. Sequence variation of a variable short tandem repeat at the D18S535 locus. Int J Legal Med, 1998, 111: 337 ~ 339

194. Lars J, Nlemann, Jette Frederiksen, Niels Morling. PCR typing of DNA fragments of the two short tandem repeat (STR) systems up stream of the human myelin basic protein (MBP) gene in Danes and Greenland Eskimos. Forensic Science International, 1996, 78: 139 ~ 155

195. Li J et al. Single nucleotide polymorphisn determination using primer extension and time - of - flight mass spectrometry. Electrophoresis, 1999, 20: 1258 ~ 1265

196. Livache T, Fouque B, Roget A et al. Polypyrrole DNA chip on a silicon device: example of hepatitis C virus genotyping. Analytical biochemistry, 1998, 255: 188 ~ 194

197. Loupart M L, Amour J A L, Jeffreys A J et al. 33.6 (D1S111) and pMLA J1 (D1S61) identify the same VNTR on chromosome 1. Nucleic Acid Research, 1991, 19 (17): 4801

198. Lygo J E et al. The validation of short tandem repeat (STR) loci for use in forensic casework. Int. J Leg. Med, 1994, 107 : 77 ~ 89

199. Markus B , Jorg DH. Versatile derivatisation of solid support media for covalent bonding on DNA ~ microchips. Nucleic Acids Res, 1999, 27 (9): 1970 ~ 1977

200. Maxam A, Gilbert W. Sequencing end labeled DNA with base - specific chemical cleavages. Methods in Enzymology, 1980, 65: 499

201. MCGall GH, Baron AD, Diggelmann M et al. The efficiency of light - directed synthesis of DNA arrays on flass substrates. J Am Chem Soc, 1997, 119 (22): 5081 ~ 5050

202. Michael klintshar. Validation of the STR system FES/FPS for forensic purposes in an Austrian population sample. Int J Legal Med , 1995, 108: 162 ~ 164

203. Moller A , Wiegand P, Gruschow C. Population data and forensic efficiency values for the STR systems hum VWA, humMBP and humFABP. Int J Legal Med, 1994, 106: 183 ~ 189

204. Morley JM, Bark JE, Evans CE et al. Validation of mitochondrial DNA minisequencing for forensic caseqwork. Int J Legal Med, 1999, 112: 22441 ~ 248

205. Morley JM, Bark JE, Evans CE et al. Validation of mitochondrial DNA minisequencing for

forensic casework. Int. J Legal Med, 1999, 112: 22441 ~ 248

206. Ngo KY et al A. DNA probe detecting multiple haplotypes of the human Y chromosome. Am J Hum Genet , 1986, 38: 407 ~ 418

207. Nu En Huang, James Schumm, Bruce Budowle. Chinese population data on three tetramericshort tandem repeat loci—HUMTH01, TPOX and CSF1PO—derived using multiplex PCR and manual typing. Forensic Science International, 1995, 71: 131 ~ 136

208. Nu En Huang, Ranajit chakaborty, Bruce Budowle. D1S80 allele frequencies in a Chinese population, Int J leg Med, 1994, 107: 118 ~ 120

209. Ohnishi Y et al. Identification of 187 single nucleotide polymorphisms (SNPs) among 41 candidate genes for sichemic heart disease in the Japanese population. Hum Genet, 2000, 106: 288 ~ 292

210. Pascali VL, Dobosz M, Brinkmann B. Coordinating Y ~ chromosomal STR research for the courts. International Journal of Legal Medicine, 1998, 112 (1): 1

211. Patrick OB, David Botstein. Exploring the word of genome with DNA microarrays. Nature Genetics supplement, 1999, 21 : 33 ~ 37

212. Radtkey R et al. Rapid, high fidelity analysis of simple sequence repeats on an electronically active DNAmicrochip. Nucleic Acids Research, 2000, 28: 17

213. Roewer L et al. Simple repeat sequences on the human Y chromosome are equally polymorphic as their autosomal counterparts. Hum Genet, 1992, 89: 389 ~ 394

214. Roger Poltl, Christine Luckenbach, Hors Ritter. The short tandem repeat locus D3S1359. Forensic Science International, 1998, 95: 163 ~ 168

215. Rolf B, Meyer E, Brinkmann B . Polymorphism at the tetranucleotide repeat locus DY389 in 10 populations reveals strong geographic clustering. Eur J Hum Genet, 1998, 6: 583 ~ 588

216. Ronald G et al. Rapid determination of single base mismatch mutations in DNA hybrids by direct electric field control. Proc. Natl Acad. Sci. USA, 1997, 94: 1119 ~ 1123

217. Ruyter D, Verstege AJ, Poel JJ et al. Five porcine polymorphisc microsatellite markers. Animal Genetics, 1994, 25: 53

218. Sanger F S et al: DNA squencing with chain - terminating inhibitors. Proc. Natl. Acad. Sci., 1977, 74: 5463

219. Savolainen P, Lundeberg J. Forensic evidence based on mtDNA from dog and wolf hairs. J Forensic Science, 1999, 44 (1): 77 ~ 81

220. Schacker U, Schneider PM, Holtkamp BE et al: Isolation of The DNA Minisatellite Probe MZ1.3 and Its Application to DNA 'Fingerprinting' Analysis. Forensic Science International, 1990, 44: 209 ~ 224

221. Schmaizing D et al. Recent developments in DNA sequencing by cappillary and microdevice electrophoresis. Electrophoresis, 1999, 20: 3066 ~ 3077

222. Schmalzing D et al. DNA typing in third seconds with a microfabricdted device.

Proc. Natl. Acad. Sci. USA, 1997, 94: 10273 ~ 10278

223. Second DNA recommendations 1991 report concerning recommendations of the DNA

224. Shewale JG, Lorino RA, Richey SL et al. Occurrence of vaiant alleles in a human population. The Journal of the Association of Genetic Technologists, 2000, 26 (4): 165 ~ 167

225. Shigeta, Yamamoto Y, Doi Y et al. Polymorphism of the D12S391 microsatellite in a Japanese population sample. Foensic Science International, 1999, 99: 61 ~ 66

226. Soteriou B, Fisher R A, Khan I M et al. Conserved gene sequences for species identification: PCR analysis of the 3′ UTR of the SON gene distinguishes human and other mammalian DNAs. Forensic Science International, 1995, 73: 171 ~ 181

227. Sparkes R et al. The validation of a 7 ~ locus multiplex STR test for use in forensic casework: (I) Mixtures, aging, degradation and species studies. Int. J Legal Med, 1996, 109: 186 ~ 194

228. Sperling FH, Anderson GS, Hickey DA. A DNA – based approach to the identification of insect species used for postmorten interval estimation. J Forensic Sci., 1994, 39 (2): 418 ~ 427

229. Spotsky R, Lipshutz RJ. Mapping genomic libaray clones using oligonucleotide arrays. Genomics, 1996, 33: 445 ~ 456

230. Technical Working Group on DNA Analysis Methods. Guidelines for a quality assurance program for DNA analysis. Crime Lab Dig, 1995, 22 (2): 21 ~ 43

231. Tsui P, Wong D. Allele frequencies of four VNTR loci in the Chinese population in Hong Kong. Froensic Science International, 1996, 79: 175 ~ 185

232. Urquhart A et al. Variation in short tandem repeat sequences ~ a survey of twelve microsatellte loci for the use as forensic identification markers. Int J Legal Med, 1994, 107: 13 ~ 20

233. Wang W, Kishida T, Fukuda M et al. The Y ~ 27H39 polymorphism in a Japannese population. Int J of Legal Med., 1996, 109 (4): 157 ~ 158

234. Weber JL, Wong C. Mutation of human short tandem repeats. Hum Mol Genet, 1993, 2: 1123 ~ 1128

235. Wiegand P, Budowle B, Rand S et al. Forensic validation of the STR systems SE33 and TC11. Int J Leg Med, 1993, 105: 315 ~ 320

236. Wiegand P, Schneider HR, Schurenkamp M. Tetranucleotide STR system D8S1132: sequencing data and population genetic comparisons. Int J Leg Med, 1998, 111: 180 ~ 182

237. Yue Gh, Beeckmann P, Bartenschlager H et al. Rapid and precise genotyping of porcine microsatellites. Electrophoresis, 1999, 20: 3358 ~ 3363

238. Zehner R, Zimmermann S, Mebs D. RFLP and sequence analysis of the cyochrome b gene of selected animals and man : methodology and forensic application. Int Legal Med, 1998, 111: 323 ~ 327